國立臺灣大學醫學院微生物學研究所
博士論文

Graduate Institute of Microbiology

College of Medicine

National Taiwan University

Doctoral Dissertation

人類組蛋白甲基轉化酶 SMYD3 在 DNA 修復功能上

扮演的角色探討

To Investigate the Function of Human Histone

Methyltransferase SMYD3 in DNA Repair

陳韻茹

Yun-Ju Chen

指導教授：鄧述諄 博士

Advisor: Shu-Chun Teng, Ph.D.

中華民國 106 年 6 月

June 2017

口試委員審定書

國立臺灣大學（碩）博士學位論文
口試委員會審定書

中文題目: 人類組蛋白甲基轉化酶 SMYD3 在 DNA 修復

功能上扮演的角色探討

英文題目: To Investigate the Function of Human Histone

Methyltransferase SMYD3 in DNA Repair

　　本論文係 陳韻茹 君(學號 D99445002) 在國立台灣大學微生物學所完成之博士學位論文，於民國 106 年 6 月 7 日承下列考試委員審查通過及口試及格，特此證明

口試委員：

＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿（簽名）
（指導教授）

＿＿＿＿＿＿＿＿＿＿　　＿＿＿＿＿＿＿＿＿＿

＿＿＿＿＿＿＿＿＿＿　　＿＿＿＿＿＿＿＿＿＿

系主任、所長　＿＿＿＿＿＿＿＿＿＿（簽名）

i

誌謝

　　首先誠摯的感謝指導教授鄧述諄老師，提供我很好的研究環境和方向的指導。另外，感謝李財坤老師、李明學老師、李財坤老師、曾秀如老師、冀宏源老師、吳國瑞老師、阮麗容老師、郭文宏醫師及林明杰醫師在口試委員會中的寶貴建議，及提供實驗材料，讓本論文更趨完整。

　　特別感謝振輝學長帶我入門，教會我所有的實驗，和我同舟共濟最後一同完成本計畫，也感謝品瑀學妹的碩班研究成果，她的成果是這篇論文的基礎。孟勳、雅嵐、沛恆、家維及所有實驗室同事在實驗上的討論和日常生活的分享，是讓我長時間在實驗室也不覺得辛苦的一大娛樂。

　　最後，最要感謝的是我親愛的家人，給我滿滿的愛和支持，讓我可以一直前進。僅以此論文獻給我愛的家人，謝謝你們！！
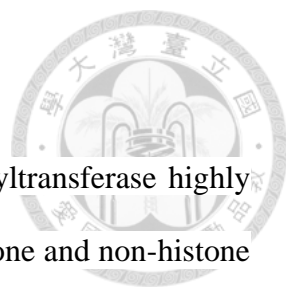
中文摘要

　　甲基轉化酶 SMYD3 高度表現在許多種類的癌症中，SMYD3 藉由甲基化組蛋白或是非組蛋白的受質來調控包括染色質重組、訊息傳導和控制細胞週期等動作。由於在癌症細胞中 SMYD3 有不正常的高度表現，且研究已指出 SMYD3 過表現會促進癌症細胞增生且連結到癌症預後不良，因此在許多種類的癌症 SMYD3 中被認為是一個預後指標。為了研究 SMYD3 的新功能，我利用 SMYD3 進行 ChIP-seq 實驗，並將其結果和先前發表的微陣列晶片資料庫做比對，實驗結果顯示一群和 DNA 損壞反應相關的基因會受到調控。進一步研究指出 SMYD3 會調控另一個導致癌症的路徑，DNA 修復。此調控是藉由促進許多參與在同源重組的基因之表現。缺乏 SMYD3 的細胞表現出對於 DNA 損壞壓力的高度敏感，DNA 斷裂及染色體重組的程度增加、修復蛋白聚集的情形降低，這些都導致同源重組功能的損害。若是在缺乏 SMYD3 的細胞中外源性的表現 SMYD3 則可以彌補這種缺陷。此外，這個對同源重組相關基因的調控是藉由在其啟動子甲基化 H3K4。這些實驗結果顯示，SMYD3 除了經由已知的機制促進癌症生成，也透過調控同源重組蛋白質貢獻於維持基因體的完整性。

關鍵字：SMYD3、組蛋白甲基化、同源重組、DNA 修復、癌症

Abstract.

SET and MYND domain containing-3 (SMYD3) is a methyltransferase highly expressed in many types of cancer. SMYD3 methylates various histone and non-histone targets to regulate distinct roles in chromatin remodeling, signal transduction and cell cycle control. SMYD3 is linked to increased cell proliferation and poor prognosis in human cancers. Due to its abnormal expression in tumors, SMYD3 is considered as prognostic markers in various cancers. To explore novel functions of SMYD3, we performed ChIP-seq experiments and compared the results with previously published microarray data. A group of DNA damage response genes were called. Further study showed that SMYD3 modulates another hallmark of cancer, DNA repair, by stimulating transcription of genes involved in multiple steps of homologous recombination (HR). Deficiency of SMYD3 induces DNA-damage hypersensitivity, increases levels of DNA breaks and chromosomal rearrangement, decreases levels of repair foci, and leads to impairment of homologous recombination. The complementation of exogenous SMYD3 is able to restore the defects caused by SMYD3 depletion. Moreover, the regulation of homologous recombination-related genes is via the methylation of H3K4 at the target gene promoters. These data imply that, besides its reported oncogenic abilities, SMYD3 may maintain genome integrity by ensuring expression levels of HR proteins to cope with the high demand of restart of stalled replication forks in cancers.

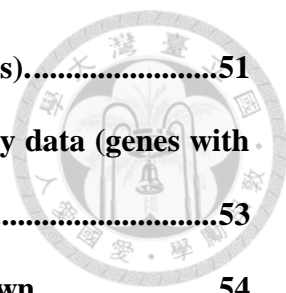Keywords: SMYD3, histone methylation, homologous recombination, DNA repair, cancer

Table of Contents.

Chapter 1. Introduction.

SMYD3 and its biological functions.

SET and MYND domain containg-3 is a member of the lysine methyltransferase family proteins, which requires S-adenosyl methionine (1) as a cofactor to methylate its substrates (2). Methylation of specific lysine residues serves as a post-translational epigenetic modification (PTM) that controls the expression of genes. It can function either by directly influence the organization of chromatin through altering histone-DNA and histone-histone interactions, or by cooperating with effector proteins which are referred to as readers of PTMs (3).

SMYD3 contains a SET domain that is divided into two segments by a MYND domain, followed by a cysteine-rich post-SET domain, and an extended C-terminal domain (CTD) (4). The MYND domain is a zinc finger motif that is important for protein-protein interaction (5). The SET domain is a conserved catalytic unit for lysine methylation that is found in nearly all protein lysine methyltransferases (6), and deletion of either the NHSC or GEEL motifs within the SET domain abolished its activity (7). The CTD domain could form a cap to bind substrates effectively and selectively (8).

Genetic mouse models for SMYD3 depletion provide hints for investigating its functions. *Smyd3* knockout mice develop normally and are fertile (9,10). However, *Smyd3* knockout mouse embryos show defects in the growth *in vitro* and a reduction in the number of viable offspring, which may be due to the suppression of the pluripotency genes, such as *Oct4*, *Nanog* and *Sox2* (11). In addition, using mouse models for pancreatic ductal adenocarcinoma and lung adenocarcinoma, the abrogating SMYD3 catalytic activity inhibits tumor development in response to oncogenic Ras (10).

1

SMYD3 is highly expressed in colorectal carcinomas, hepatocellular carcinomas, pancreatic cancer, prostate cancer, and breast cancer (7,12,13). Its high expression correlates with an aberrant pattern of histone modifications which causes abnormal gene expression (Appendix Fig. 1). SMYD3 regulates gene transcription through methylating histone substrates, including H3K4me2/3 (14), H4K20me2/3 (4), H4K5me1/2/3 (9) and H2A.ZK101 (15). For example, SMYD3 methylates H2A.Z to activate cyclin A1 expression and drive cancer proliferation (15), and H3K4 to upregulate MMP9 (16) and hTERT expression (17). Moreover, SMYD3 modifies non-histone proteins VEGFR and MAP3K2 to promote metastasis (18) and Ras/Raf/MEK/ERK signaling (10) in cancer development, respectively. SMYD3 mainly locates in the cytoplasm at G0/G1 phases and moves to the nucleus at S/G2 phases (19). Therefore, the function of SMYD3 may depend on its location.

The involvement of SMYD3 in human pathology is not restricted to cancer. SMYD3 controls the proper development of skeletal muscle via the regulation of c-met and myostatin (20), a critical negative regulator of cell differentiation and muscle mass (21). SMYD3-depleted cell culture and the animal model show decreased expression of c-met and myostatin, resulting in protection against glucocorticoid-induced muscle atrophy. Notably, SMYD3 are more abundantly expressed in skeletal muscle compared with other tissues (19), which enables it an attractive target for the treatment of muscle disease with increased specificity.
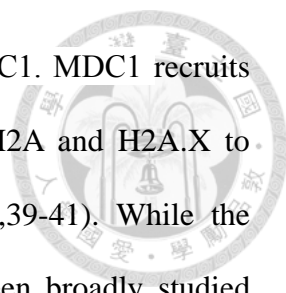
DNA double-strand breaks repair pathways.

Cells inevitably encounter the challenge of chromosomal double-strand breaks (DSBs) during their lifetime. The oxidative byproducts of the normal metabolic process and exogenous factors such as chemical agents or ionizing radiation (IR) constantly

2

threaten the integrity of our genome. Unrepaired or misrepaired DNA lesions can lead to genome instability, which is a hallmark of cancer (22,23). Two major pathways, homologous recombination and non-homologous end joining (NHEJ), are responsible for repairing these breaks.

HR occurs predominantly at S and G2 phases when a sister chromatid is accessible (24). The repair is initiated by a resection process, which includes MRE11-RAD50-NBS1 (MRN) end sensing complex, CtIP endonuclease (25), EXO1 exonuclease (26), and BLM helicase (27), to remove oligonucleotides from each side of the DSB and expose single-stranded DNA (ssDNA) tails for forming RAD51-ssDNA filaments with the help of BRCA2 (28). Working in concert with RAD51-ssDNA filaments, RAD54B, a DNA-dependent ATPase, drives the search for a homologous template and strand invasion (29,30), which leads to accurate repair. Classical NHEJ (C-NHEJ) occurs throughout the cell cycle but predominately at G1 phase. During C-NHEJ, cells utilize Ku70/Ku80 heterodimer and DNA-dependent protein kinase (DNA-PK) to recognize and ligate DSB ends via little (less than ten base pairs) or no homology between the joined ends, which is, therefore, an error-prone pathway (31,32). Besides, alternative end-joining pathways, such as microhomology-mediated end joining (MMEJ), do not use Ku- and DNA-PK. Initial resection produces relatively longer stretches of microhomology (5-25 base pairs), and subsequent flap trimming and end-joining often create the final mutagenic MMEJ repair products (33).

DSB repair is facilitated through chromatin modifications to open the compact barriers and improve the accessibility of repair proteins (34). For example, the rapid phosphorylation of H2A.X at S139 in mammals (forming γH2A.X) by ATM within minutes at DSB sites is considered as a major hallmark of DSB recognition (35,36).

3

γH2A.X further interacts with the mammalian repair mediator MDC1. MDC1 recruits RNF8 (37) and RNF168 (38) to catalyze K63-ubiquitilation on H2A and H2A.X to recruit BRCA1 and 53BP1 for HR and NHEJ, respectively (37,39-41). While the posttranslational modifications of proteins in DSB repair have been broadly studied (42,43), the evidence of transcriptional regulation of DSB repair proteins is comparatively scarce.

SMYD3 is involved in homologous recombination.
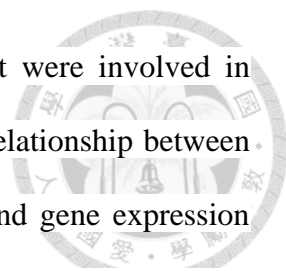
Previous studies have focused on the ability of how SMYD3 stimulates cell proliferation and metastasis. Here, we identify a new role of SMYD3 in regulating HR repair. After IR treatment, SMYD3-depleted cells showed a delay in the removal of γH2A.X foci, and are more vulnerable to the IR stress. The ratio of DNA breaks and the gross chromosomal rearrangement are increased in SMYD3-deficient cells post IR treatment. Moreover, inhibition of SMYD3 directly blunts HR efficiency by downregulating the expression of HR-related genes. Additionally, SMYD3 knockdown leads to decreased methylation of H3K4 and recruitment of RNA polymerase II (RNAPII) at the target gene promoters. These data reveal that SMYD3 maintains genome stability by ensuring normal expression levels of HR repair proteins.

4

Chapter 2. Results.

Chromatin immunoprecipitation followed by high-throughput sequencing (ChIP-seq) analysis reveals a role of SMYD3 in DNA repair.
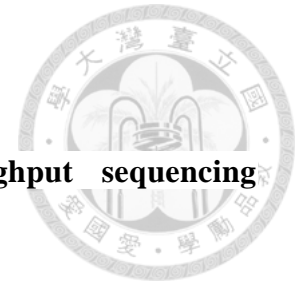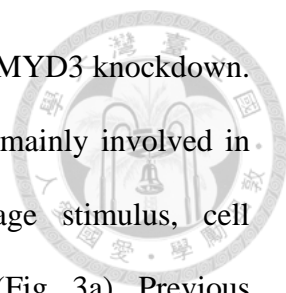
In previous work (15), we have identified that cyclin A1 is a downstream target of SMYD3 and H2A.Z.1. In order to characterize at a genomic scale the unexplored aspects of the relationship among SMYD3, H2A.Z.1, and their cellular targets, we performed SMYD3 and H2A.Z.1K101me2 ChIP-seq experiments. We performed three times of paired-end sequencing for each sample to get a total of 30 million reads. The output results were pooled together for further analysis. The signals of SMYD3 and H2A.Z.1K101me2 were normalized to input signals (Fig. 1a). We compared the distribution map of about 20 thousand genes with reported H2A.Z and H2A.Zac (acetyl K4 + K7 + K11) ChIP-seq database (GEO accession number GSM1059388) (44). Results showed that SMYD3, H2A.Z, and H2A.Z PTM proteins did not share similar tendency of distribution, in which SMYD3 was recruited more at the promoter region (-5000 bp to -1000 bp) than at coding sequences, whereas H2A.Z and its PTM proteins were more centered around the transcription start site (TSS). Compared to H2A.Z and H2A.Zac, which showed equal distribution around TSS, H2A.ZK101me2 signal levels were higher at the promoter region (-1250 bp to TSS) than at the downstream of the TSS. Next, we examined unique genes identified from the ChIP-seq data. Using cutoff parameters of $P < 0.001$ and distance to TSS (-5000 bp to +200 bp), SMYD3 and H2A.Z.1K101me2 were associated with 2406 and 3029 genes, respectively (Fig. 1a and 1b). Gene ontology analysis showed that SMYD3 regulated genes involved in regulation of transcription factor activity, DNA binding, enzyme activity and

5

nucleocytoplasmic transport. H2A.ZK101me2 regulated genes that were involved in protein transport and cell cycle. Furthermore, we investigated the relationship between the enrichment of SMYD3 and H2A.Z.1K101me2 on chromatin and gene expression levels. We compared their occupancy and the gene expression data from our previous report (15). We then separated genes into four classes based on their increase or decrease in expression, namely they were downregulated or upregulated in shLuc versus shSMYD3 cells (Table 1 and 2), and in H2A.Z.1^{WT} versus H2A.Z.1^{K101Q} in endogenous H2A.Z knockdown MCF7 cells (Table 3 and 4). Since SMYD3 is a transcriptional co-activator (45), we focused on genes that were downregulated in the shSMYD3 cell (Fig. 2a and Table 1). Gene ontology (GO) analysis showed that 5 genes were involved in DNA repair, including *EXO1*, *FANCI*, *MDC1*, *POLQ*, and *RAD54B*, and 4 genes were related to M phase, including *ANLN*, *EXO1*, *OIP5*, and *RAD54B* (Table 5). To verify the accuracy of the microarray data, qRT-PCR was performed for 5 DNA repair genes, and all genes exhibited similar fold as originally identified in the microarray analysis (Fig. 2b). We then went back to review the microarray data and found that GO analysis suggested a role of SMYD3 in response to DNA damage stimulus (Fig. 3a). Therefore, we aimed to study whether SMYD3 is involved in DNA repair.

Microarray data analysis identifies SMYD3-regulated expression of DNA repair machinery.

Increased expression of SMYD3 can promote cancer proliferation (7) and metastasis (16). To explore additional and novel roles of SMYD3 in biological processes, we analyzed our previously conducted whole-genome microarray data of RNAs isolated from shLuc vs. shSMYD3 MCF7 cells (GEO accession number GSE58048), in which a lentivirus shRNA infection system was used for stable

6

knockdown of SMYD3 (15). 449 genes were downregulated upon SMYD3 knockdown. The gene ontology (GO) analysis indicated that these genes were mainly involved in cell cycle, DNA metabolic process, response to DNA damage stimulus, cell proliferation and macromolecular complex subunit organization (Fig. 3a). Previous reports provided evidence that SMYD3-dependent histone methylations are essential for cell cycle and cell proliferation. Intriguingly, SMYD3 is associated with DNA damage response (DDR) in the top three categories of GO analysis. Since SMYD3 has not been linked to DDR or DNA repair, the mechanism was further analyzed.

We first investigated whether SMYD3 knockdown cells are more vulnerable to DNA damage stress such as IR. To examine the repair rate following IR, the formation of γH2A.X foci were used as a marker for DNA damage. Following exposure to 1.67 Gy of IR, SMYD3 knockdown cells significantly delayed the removal of γH2A.X foci at 48 and 72 hr compared to the shLuc controls (Fig. 3b). Conversely, exogenous expression of SMYD3, but not the catalytic dead mutant SMYD3^{Y239F} proteins, in the SMYD3 knockdown cells significantly restored the defects at 24 and 72 hr compared to that in the shSMYD3 with the expression of the vector control (Fig. 3c). The clonogenic survival following exposure to IR was further examined, and knockdown of SMYD3 led to impeded formation of colonies (Fig. 4a). Similar to that in MCF7 cells, shSMYD3 MDA-MB-231 and AU565 cells were more vulnerable to IR stress (Fig. 4b and 4c).
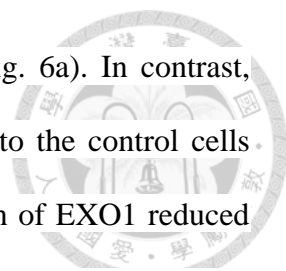
We next analyzed the effect of IR on SMYD3's cellular location. The majority of SMYD3 is located in the cytoplasm (10), and we wondered whether it would translocate into the nucleus upon DNA damage insults. Cells were exposed to increasing dosages of IR and assayed for the translocation of SMYD3 protein to the nucleus at 1 and 3 hours.

7

The distribution of SMYD3 did not show any noticeable difference after IR treatment (Appendix Fig. 2a). Furthermore, the gene and protein expression of SMYD3 were not augmented after IR treatment (Appendix Fig. 2b). These results suggest that DNA damage does not modulate SMYD3 expression and location.

We further examined whether SMYD3 affects genome integrity in cells. To determine if the loss of SMYD3 is associated with increased DNA damage, we performed a single-cell gel electrophoresis (comet) assay. The comet assay revealed increases of damage rate and tail moment in shSMYD3 compared to those in shLuc cells at 72 h (Fig. 5). We also investigated micronuclei formation, a well-established indicator of genome instability (46,47), which occurs through the aberrant segregation of chromosomes or acentric chromosomal fragments. Compared to the shLuc control, knockdown of SMYD3 caused an increase in IR-induced micronuclei at 72h (Appendix Fig. 3). Taken together, these data suggest a role of SMYD3 in DNA repair mechanism.

SMYD3 mediates the HR pathway.

Since mammalian DSB repair was achieved mainly by two mechanisms, HR, and NHEJ, we wondered whether SMYD3 is involved in these pathways. We used cells with well-characterized GFP-based chromosomal reporters to detect the efficiency of HR. The reporter contains an I-SceI recognition sequence, which would be cleaved upon I-SceI expression to generate a DSB. DSB repair by HR using the direct repeat within the reporter cassette as a template results in an intact GFP gene. The repair efficiency was then quantified by flow cytometry. For the plasmid-based end-joining assay, a linearized plasmid harboring a luciferase reporter gene was used. Repair efficiency was measured by the luciferase activities of linearized reporter constructs compared with that of the intact plasmid. Results demonstrated that SMYD3 knockdown significantly

8
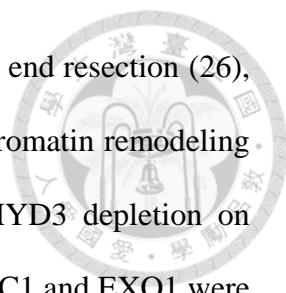
hampered HR repair by 55-70% compared to the control cells (Fig. 6a). In contrast, SMYD3 knockdown did not change the NHEJ activity compared to the control cells (Fig. 6b). As the controls for the HR and NHEJ assays, knockdown of EXO1 reduced the HR activity and knockdown of Ku70 impaired the NHEJ activity by 56-73% and 47-50%, respectively (Fig. 6a and 6b). We also examined whether SMYD3 was required by MMEJ using a plasmid-based MMEJ assay and found that SMYD3 did not display any effect on the efficiency of MMEJ, while knockdown of the control, POLQ, reduced MMEJ activity by 49-58% (Fig. 6c). The knockdown effectiveness of each cell lines used was confirmed by qRT-PCR (Fig. 7). Moreover, the exogenous expression of SMYD3, but not the SMYD3^{Y239F} proteins, restored the HR activity of the shSMYD3 cells (Fig. 8). These results identify a role of SMYD3 in HR repair.

SMYD3 knockdown downregulates HR gene expressions.

To understand the exact role of SMYD3 in HR repair, we analyzed microarray-identified genes that were related to DDR and found that 32 of them are involved in DNA repair, including genes that are in response to oxidative stress, base excision repair, cell cycle checkpoint, interstrand cross-links, HR, NHEJ and MMEJ (Table 6). Among these 32 genes, 13 genes are implicated in the HR pathway (Table 7). These genes range from the early step of DNA damage mediators, kinase transducer to downstream effectors that execute error-free repair process (48). We performed qRT-PCR analysis to confirm their expressions and found that all genes exhibited similar fold differences in mRNA expression as initially identified in the microarray analysis (Fig. 9).

Among these genes, MDC1 plays the earliest role in HR repair (49) and participates in the initial recruitment of BRCA1 (50-54) to promote DNA end resection

9

for HR (55). Moreover, EXO1 is the major exonuclease for efficient end resection (26), and RAD54B is a DNA-dependent ATPase required for efficient chromatin remodeling during strand invasion (29,30). We checked the influence of SMYD3 depletion on MDC1, EXO1, and RAD54B. Consistently, the protein levels of MDC1 and EXO1 were significantly reduced after inhibition of SMYD3. Besides, a marginal reduction of the RAD54B protein was observed (Appendix Fig. 4).

To gain further insight into how SMYD3 regulates HR activity, we investigated the significance of SMYD3 on the formation of MDC1 foci. After IR treatment, MDC1 foci were diminished in SMYD3-depleted cells at 1 to 4 hrs. And these MDC1 foci were disappeared at 24 to 72 hrs, even in the shLuc cells (Fig. 10a). Moreover, after IR treatment, the formation of BRCA1 foci in SMYD3 knockdown cells was impaired as well (Fig. 10b). In contrast, lack of SMYD3 did not affect the assembly of 53BP1 foci (Appendix Fig. 5). These data indicated that SMYD3 deficiency weakens HR partly through downregulation of MDC1, and thereby compromising the recruitment of BRCA1 at DSBs.

SMYD3 knockdown impaired the formation of BRCA1 foci. We further investigate whether it was resulted from the change of mRNA and protein expression of BRCA1 *per se* or affected by the reduction of MDC1. We observed that the mRNA and protein level of BRCA1 diminished in SMYD3-depleted cells (Appendix Fig. 6A and 6B). Intriguingly, our microarray data showed that the expression of *BARD1* (BRCA1-Associated RING Domain-1) decreased in SMYD3-depleted cells (Appendix Fig. 6A). BARD1 is vital in the rapid relocation of BRCA1 to DNA damage sites, and the interaction with BARD1 increases the E3 ligase activity and the stability of BRCA1 (56). We observed that MDC1-depleted cells showed decreased BRCA1 foci/protein

10
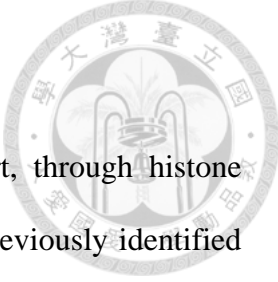
expression as well (Appendix Fig. 6C and 6D). Therefore, the reduction of BRCA1 foci/protein in SMYD3-depleted cells might due to the synergistic effects caused by BARD1 and MDC1.
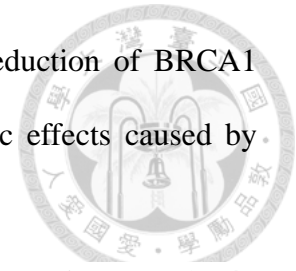
SMYD3 controls the expression of HR genes through methylating histone H3K4.

SMYD3 was initially reported to methylate histone H3K4 to modulate the accessibility of chromatin architecture and to form a complex with RNA polymerase II (RNAPII) and RNA helicase HELZ to drive its target genes (7). To explore the epigenetic regulation of SMYD3 on these HR genes, we examined the recruitments of SMYD3, H3K4me3, and RNAPII pSer5, which is a required RNAPII phosphorylation for the transcriptional initiation (57), to the *MDC1*, *EXO1,* and *RAD54B* promoter regions. The putative TATA box (region TA) was retrieved from GPMiner (58). ChIP experiments showed a direct binding of SMYD3 to the *MDC1* promoter (region S3) at ~500 bp upstream of the transcription start site (TSS) (Fig. 11a). SMYD3 preferred to enrich at region S3 (blue bar) than region TA (red bar) (Fig. 11b). In contrast, RNAPII pSer5 (Fig. 11c) and histone H3K4me3 (Fig. 11d) exhibited greater binding at region TA than region S3. Furthermore, knockdown of SMYD3 led to significant decreases of SMYD3, H3K4me3, H3 and RNAPII pSer5 at various degrees at both regions S3 and TA (Fig. 11b-e). Also, the enrichment of H3K4me3-modified histones (H3K4me3/H3) declined significantly in SMYD3 knockdown cells (Fig. 11f). A similar tendency was observed in the promoter regions of *EXO1* and *RAD54B* (Fig. 12). These results suggest that SMYD3 may trigger HR gene expression by directly binding to the promoters to create an active histone mark for transcription.

11

Chapter 3. Discussion.

SMYD3 is a transcriptional regulator that functions, in part, through histone methylation to control the expression of target genes. Apart from previously identified targets *WNT10B* (12), *hTERT* (17), *CCNA1* (15), *NKX2.8* (7), *MMP9* (59), and *c-MET* (60), we show here that SMYD3 facilitates expression of several genes involved in the whole process of HR repair. Deficiency of SMYD3 hampers the expression of these HR associated genes, induces DNA-damage hypersensitivity, causes genomic instability, decreases the levels of MDC1 and BRCA1 foci, and leads to impairment of HR-mediated DSB repair. Furthermore, generation of the active transcription mark H3K4me3 and phosphorylation of RNAPII c-terminus Ser5 at the *MDC1*, *EXO1,* and *RAD54B* promoters are SMYD3-dependent. These findings provide insights into how SMYD3 functions as an oncogene besides its abilities in promoting cell cycle, cell proliferation, and metastasis. The newly identified role of SMYD3 in HR repair proves that SMYD3's function is crucial in maintaining genome stability (Fig. 13). Consistent with recent reports, highly proliferative cells, such as cancer cells, rely profoundly on HR-mediated DSB repair to restart the stalled replication forks at S phase (61,62).

The activity of SMYD3 is closely related to the cell cycle progression (15,63). Deficiency of SMYD3 leads to G1-phase cell cycle arrest in breast cancer cells (16,63) and S-phase arrest in prostate cancer cells (64). Therefore, it is likely that slower removal of IR-induced γH2A.X foci (Fig. 3b) might also be contributed partly from the impact of SMYD3 on cell cycle. Notably, SMYD3 deficient cells could proceed the cell cycle to finish mitosis, but generate more micronuclei which represent aberrant chromosomal segregation (Appendix Fig. 3). The combinational effects of genomic

12

instability and the lack of proper DNA repair machinery in SMYD3-depleted cells may therefore lead to apoptosis (63,64).

Large-scale screens have revealed a correlation between SMYD3 and DNA repair pathways. Knockdown of SMYD3 caused differentially expressed DNA repair genes *RAD50* and *RAD51* in prostate cancer cells (65). A proteomic analysis discovered that SMYD families share some common interactors (NPM1, TOP1, GNL3, and RUVBL2) involved in DNA repair and chromatin maintenance (44), which may imply that SMYD3 not only modulates the DNA repair pathways at the transcription level but also directly interacts with DDR proteins. Moreover, mono-methylation of poly(ADP-ribose) polymerase-1 (PARP1) by SMYD2, another SMYD family methyltransferase, enhances PARP1 activity and cellular response to oxidative DNA damage (66). Therefore, we propose that SMYD3 may regulate DNA repair at both transcriptional and posttranscriptional levels.

Most histone methylation occurs at specific sites of H3 and H4, which is controlled by a large group of methyltransferases and demethylases. For example, SETD2-dependent H3K36me3 is essential for efficient end resection of HR through the recruitment of CtIP, RPA, and RAD51 (67,68). Interestingly, the levels of SETD2 and H3K36me3 are not induced after DSB damage, suggesting their pre-established role on chromatin (68,69). Similarly, the distribution and protein amount of SMYD3 are not adjusted after IR treatment (Appendix Fig. 2). These finding may imply that the amount of nuclear SMYD3 is sufficient for driving HR gene expression and that these HR proteins are constitutively expressed for taking care of sudden DNA damage. The property that SMYD3 is preferentially recruited to individual promoters may rely on both its binding sequences and its binding partners. For instance, SMYD3 cooperates

13

functionally with PC4 to activate the expression of genes that promote cell proliferation and invasion in a cancerous background. Both proteins localize to their common targets in a dependent manner, which suggests that SMYD3 relies on PC4 interaction to stabilize its binding to growth-promoting target genes and upregulate their expression (70). SMYD3 may also respond to diverse general and tissue specific signals to control its transcriptional activation and subcellular localization (13). Further investigation is required to clarify this specificity.

Chapter 4. Materials and methods.

Cell lines, plasmid construction, and gene knockdown

The breast cancer cell lines MCF7 and MDA-MB-231 were maintained in DMEM, and AU565 were maintained in RPMI. These mediums were supplemented with 10% FBS, non-essential amino acids (HyClone), sodium pyruvate (HyClone), and antibiotics (HyClone). HEK293T cells were co-transfected with packaging plasmid (pCMV-Δ8.91), envelope (pMDG), and hairpin pLKO-RNAi vectors (National RNAi Core Facility, Institute of Molecular Biology/Genomic Research Centre, Academia Sinica, Taiwan) for virus packaging. The specific oligo sequences of shRNA are listed in Table 8. Virus-containing supernatants were collected at 48 hr post-transfection. Cells were treated with virus plus medium containing polybrene (8 μg/ml) for 16 hr. The infected cells were selected with puromycin (0.5 μg/ml). Plasmids expressing SMYD3^{WT} and methyltransferase-dead SMYD3^{Y239F} were constructed as previously described (15).

ChIP assay.
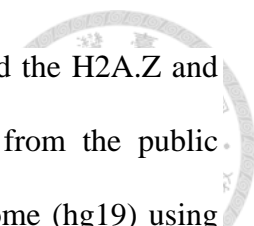
ChIP assays were performed as described (15). Complexes were immunoprecipitated overnight with 2 μg of antibodies specific for SMYD3 (GTX121945, GeneTex), rabbit IgG (GTX35035, GeneTex), H3 (ab1791, Abcam, Cambridge, MA, USA), H3K4me3 (ab10158, Abcam), and RNA polymerase II CTD repeat YSPTSPS (phosphor Ser5) (ab5131, Abcam). Input samples were processed in parallel. Antibody/protein complexes were collected by 40 μl of protein G-coupled Sepharose beads (GE Healthcare Bio-Sciences, Pittsburgh, PA, USA) and washed as follows: once with Tris/EDTA-150 mM NaCl, twice with Tris/EDTA-500 mM NaCl, and once with PBS. Immune complexes were eluted with 1% SDS and TE buffer. After

15

decrosslinking, DNA was purified using a PCR cleanup kit (Qiagen) and analyzed by qRT-PCR. The results were expressed as the percentage of the initial inputs. The primer sets used for the ChIP assay are listed in Table 9.
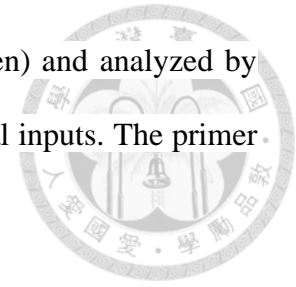
ChIP-seq assay and data analysis.

ChIP DNA was obtained as described above. For library construction, 3 samples were prepared, including input DNA (3 μg), SMYD3-ChIP DNA (100 ng) and H2A.ZK101me2 (45 ng). The samples were adjusted to equivalent concentration for later steps. Illumina adaptor sequences and barcodes were added in a subsequent limited PCR reaction, and the libraries were sequenced with an Illumina MiSeq platform (Illumina, San Diego, CA, USA) at the Technology Commons in College of Life Science and Center for Systems Biology (National Taiwan University, Taiwan). Three times of paired-end sequencing for each sample to get a total of 30 million reads were conducted at a read length of 150 bp. Raw sequences were obtained using the Illumina GA Pipeline software CASAVA v1.8. The sequences generated were subjected to a filtering process to obtain qualified reads. ConDeTri was implemented to trim or remove reads according to their quality score (71). The read data were aligned to the genomic sequences retrieved from the UCSC database using BWA, and these results were imported into the Partek Genomics Suite (Partek Inc., St. Louis, MO) to identify peaks in the ChIP-seq data using a standard ChIP-seq workflow (72). A zero-truncated negative binomial model was fitted to the sequencing data to identify the enriched regions as peaks (Partek ChIP-Seq white paper: www.partek.com/Tutorials/microarray/User_Guides/ChIPSeqPeakDetection.pdf). The p-value from the Mann-Whitney U test and the binomial p-value ($p < 0.001$) were used
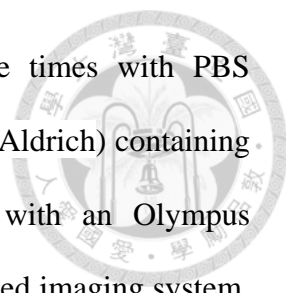
16

to remove any false peaks. In the ChIP-seq data analysis, we collected the H2A.Z and acetylated-H2A.Z datasets (GEO accession number GSM1059388) from the public domain, and these sequences were mapped to the UCSC human genome (hg19) using the Bowtie aligner (according to the following parameters: -a -m 1 -v 3) (73). We then applied the MACS peak caller to the mapped reads to find the significant peaks ($p <$ 0.001), and the peaks were annotated with the corresponding gene information using ChIPpeakAnno (74). Only the identified peaks in the promoter region (5 kb upstream and 0.2 kb downstream of the transcription start site) were captured, and we obtained candidate genes as those with promoter regions that were enriched based on the ChIP-seq experiment.

Data analysis of gene expression microarray.

The microarray analysis of shLuc vs. shSMYD3 viruses-treated MCF7 cells was performed as previously described with the NCBI Gene Expression Omnibus (GEO) accession number GSE58048 (15). The list of down-regulated genes was further categorized by the DAVID v6.8 Gene Ontology program (75,76).

Cell fixation and immunofluorescence assays.

Cells were seeded on glass coverslips coated with poly-L-lysine (Sigma-Aldrich, St. Louis, MO, USA) and allowed to attach for 48 hr followed by 1.67 Gy IR treatment (IBL 637, CIS Bio International, Gif-sur-Yvette, France). After washed with phosphate buffered saline (PBS), cells were fixed in 4% paraformaldehyde in PBS for 10 min at room temperature. The fixed cells were then permeabilized with 0.1% Triton X-100 in PBS for 10 min. After 30 min of blocking with 1% BSA in PBS, cells were washed in PBS and incubated with primary antibodies for 3 hr. After three times washed in PBS containing 0.05% Triton-X for 5 min, the cells were incubated with

17

secondary antibodies for 1 hr. Finally, cells were washed three times with PBS containing 0.05% Triton-X and embedded in 1 μg/ml DAPI (Sigma-Aldrich) containing mounting solution on glass slides. The cells were visualized with an Olympus fluorescence microscope. Images were captured using a Spot advanced imaging system. The primary antibodies used were γH2A.X (1:200, 05-636, Millipore-Upstate, Temecula, CA, USA), BRCA1 (1:100, sc-6954, Santa Cruz Biotechnology, CA, USA), 53BP1 (1:100, sc-22760, Santa Cruz Biotechnology) and MDC1 (1:200, A300-053A, Bethyl Laboratories, Montgomery, TX, USA).

Colony formation assay.

For the colony formation assay, control (shLuc) or knockdown (shSMYD3) cells were seeded (5,000 cells for shLuc cells, and 20,000 and 10,000 cells for shSMYD3#1 and shSMYD3#2 cells, respectively) in 6-cm dishes two days before IR (0, 1.67, 3.34, 5.01 Gy) treatment. Cells were incubated for 15 days, fixed in 4% paraformaldehyde for 5 min, washed once with PBS, stained with 0.1% crystal violet, and then washed with distilled water. The survival rate was calculated by comparing the colonies numbers with the non-irradiated cells in each group.

Nuclear/cytosol fractionation.

Approximate 1×10^6 MCF7 cells were trypsinized and washed with ice-cold PBS twice followed by lysed on ice for 10 min in 250 μl cytoplasmic lysis buffer (0.1% Triton-X, 10 mM HEPES-KOH pH7.9, 10 mM KCl, 1.5 mM $MgCl_2$, 0.34 M sucrose, 10% glycerol) containing protease inhibitor, 1 mM DTT, and 10 mM PMSF. Nuclear sediments were collected by centrifugation at 6,000 rpm for 1 min, and pellets were washed twice with 1 ml cytoplasmic lysis buffer. Nuclei were lysed in RIPA buffer (50 mM Tris-HCl, 150 mM NaCl, 1% NP-40, 0.25% sodium deoxycholate, 1% sodium

18

dodecyl sulfate (SDS), 1 mM DTT, protease inhibitor, 1 mM PMSF, 1 mM EDTA) and lysed completely by sonication. Nuclei and cytosolic extracts were then subjected to Western blot analysis.
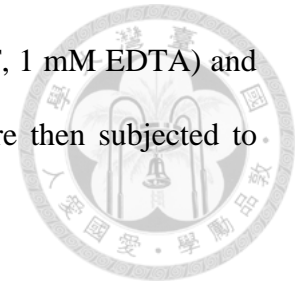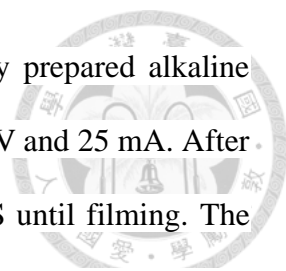
Western blot analysis.

Western blot analysis was performed as described(15). The primary *antibodies used* were SMYD3 (1:3000, GTX121945, Genetex, San Antonio, TX, USA), MDC1 (1:1000, GTX102673, GeneTex), EXO1 (1:1000, GTX109891, GeneTex), RAD54B (1:1000, GTX103291, GeneTex), Tubulin (1:5000, GTX112141, GeneTex), and nuclear matrix protein p84 (1:10000, NB100-174, Novus, Littleton, CO, USA). The quantification of protein expression was performed using ImageJ software (Image Processing and Analysis in Java). All protein expression levels were normalized against the corresponding control protein levels as indicated. Images were representatives of n ≥ 3 for each experiment.

Comet assay.

DNA strand breaks were evaluated using alkaline single cell gel electrophoresis (comet) assay following the procedure of Olive and Banath (77). In brief, 2×10^5 cells with 4-day SMYD3 knockdown were seeded in 6-well tissue-culture plates and incubate for 24 h for cell attachment. Cells were treated with 1.67 Gy IR and harvested for indicated time. Cells were washed with PBS twice and resuspended in 1 ml PBS. About 50 µL of the resuspended cells was mixed with 100 µL of low melting point agarose at 45 ℃ and spread the suspension over the well with the pipette tip. The slides were placed at 4 ℃ in the dark until gelling occurred and then immersed in pre-chilled lysis buffer at 4 ℃ for 30 min. After that, the buffer was aspirated and replaced with a pre-chilled alkaline solution for 40 min at 4 ℃. After lysis and unwinding, the slides

19

were placed in a horizontal electrophoresis tank filled with freshly prepared alkaline electrophoresis buffer. The electrophoresis was run for 15 min at 15 V and 25 mA. After electrophoresis, the slides were rinsed with PBS and placed in PBS until filming. The slides were added with 100 µL of 10000X diluted SYBR gold (Thermo Fisher Scientific Inc., Waltham, MA, USA) for 1 min for DNA staining. DNA migration was observed using fluorescence microscope at a magnification of 20X (Carl Zeiss Apo Tome, Germany). For each condition, at least 100 randomly selected cells were analyzed. The quantification of the tail moment, which is a representation of the fluorescence intensity in the tail relative to the head, was performed using CometScore (Autocomet.com).

Micronuclei counts.

For micronuclei analysis (46,47), cells were seeded and fixed as described in the immunofluorescence assays. The cells were then incubated with SYBR gold (Thermo Fisher Scientific Inc., Waltham, MA, USA) for 1 min, washed twice with PBS, and mounted for microscopy. Objects were defined as micronuclei if they were clearly separated from the nuclei, were round- to oval-shaped with distinct borders, had an area of less than a quarter of the area of a nucleus, and showed staining characteristics similar to those of nuclei.

RNA analysis and quantitative real-time polymerase chain reaction (qRT-PCR).

Total RNA was isolated using the RNeasy kit (Qiagen, Valencia, CA, USA). RNA was reverse-transcribed into first-strand cDNA using AMV reverse transcriptase (Promega, Madison, WI, USA). cDNA was amplified with KAPA SYBR Fast PCR Mix (KAPA Biosystems, Woburn, MA, USA) and subjected to analysis using a *CFX Connect* Real-Time System thermal cycler (Bio-Rad, Hercules, CA, USA). *RPL30* mRNA, which encodes the ribosomal protein L30, was used as an internal control. The

20

relative abundance of mRNA was calculated after normalization with *RPL30* mRNA using the CT equation. For verifying candidate genes in microarray data, primers were either designed based on the coding sequence of each genes using Primer3Plus(78) or directly retrieved from Origene website (http://www.origene.com/). The primers used are listed in Table 9.

HR assay.

HR efficiency was measured in MCF7/DR-GFP cells, according to the previous report (79). The MCF7 DR-GFP cells harbor GFP-based chromosomal reporters. *The* stable cells possess two differential GFP mutant genes oriented as direct repeats and separated by a drug selection marker, the puromycin N-acetyltransferase gene. Transient expression of the I-SceI enzyme produces a DSB in one of the two *GFP* mutant genes. The DSB can be repaired by HR between the two GFP mutant genes, resulting in the restoration of a functional GFP gene and the expression of GFP proteins. After knockdown of target genes for three days, cells were transfected with pCASce to express the I-SceI protein. GFP-positive cells were measured by flow cytometry (BD FACSCalibur, BD Biosciences, Miami, FL, USA) in 48 hr. Ds-Red was transfected in a parallel group as a control for transfection efficiency.

Plasmid based end-joining assay.

Plasmid end-joining assay was conducted as previously described (80). The pGL3-promoter plasmid (Promega), which harbors a luciferase reporter gene, was linearized by HindIII and confirmed by agarose gel electrophoresis. The linearized DNA was purified by gel extraction kit (Qiagen), dissolved in sterilized water, and transfected into cells after knockdown of target genes for three days. Luciferase protein was expressed when the cutting sites were repaired by end-joining. The luciferase activity

21

was assayed by Luciferase Assay System (Promega). A Renilla plasmid was co-transfected as a control.

Plasmid based MMEJ assay.

MMEJ efficiency was measured according to the previous report (81). The pSV40-MMEJ plasmid, which harbors a GFP reporter gene, was a gift from Dr. Nicolas Mermod. Plasmids were linearized by I-SceI, purified by gel extraction kit (Qiagen), and transfected into cells after knockdown of target genes for three days. GFP protein was expressed when the cutting sites were repaired by MMEJ. The pGK-GFP plasmid was transfected in parallel as a transfection efficiency control. Expression of GFP was measured by flow cytometry (FACSCalibur, BD BioSciences) after 48 hr of transfection.

Statistical analysis.

Each experiment was repeated at least three times with comparable results. Results are expressed as mean ± SD. All statistical analyses were performed using Excel 2010 (Microsoft; Redmond, WA). The *p*-values for all experiments were obtained using Student's *t* tests or one-way analysis of variance (35). A value of $P < 0.05$ was considered to be statistically significant.

Chapter 5. Figures and Figure Legends.

a

b SMYD3 peak calling process

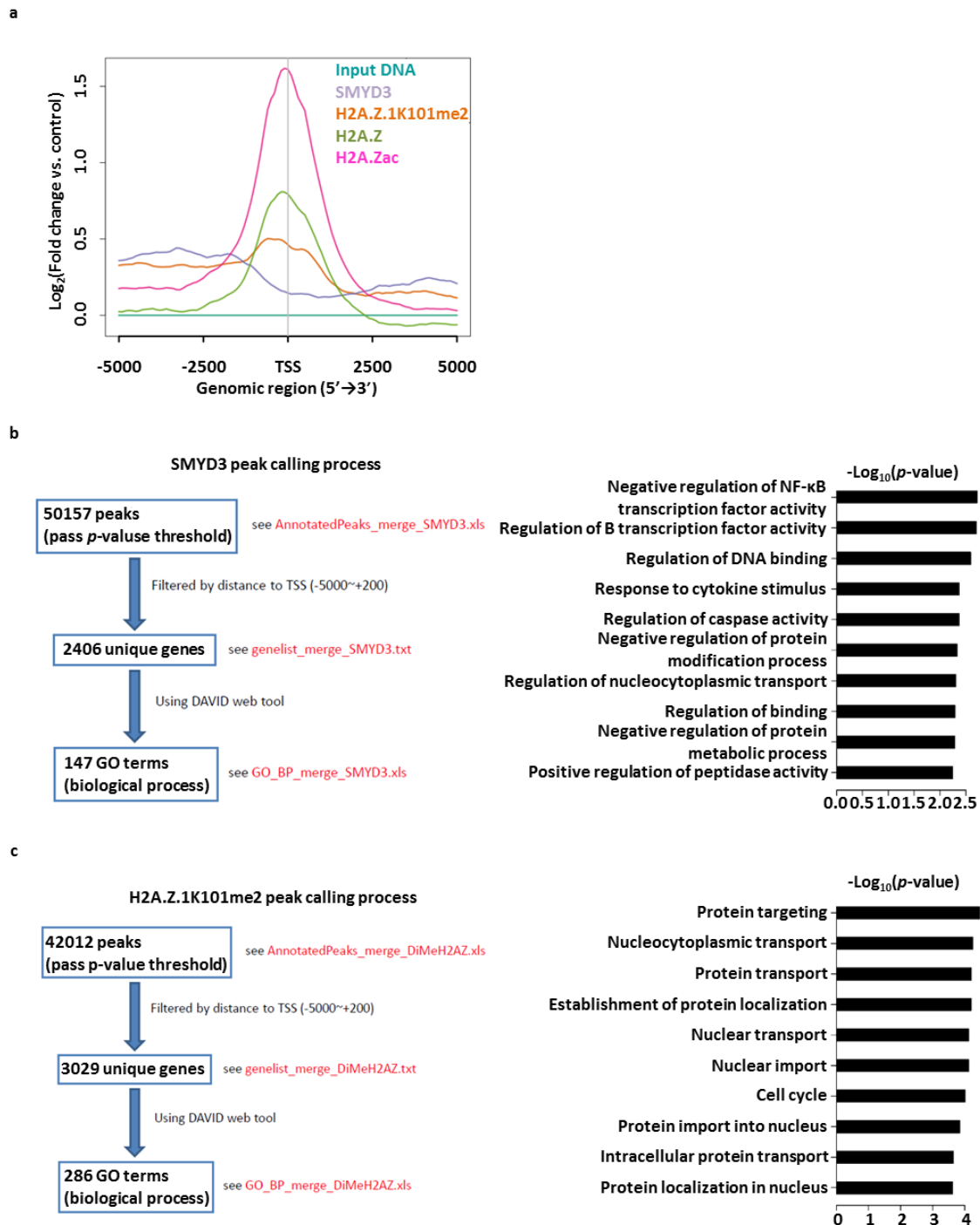**c** H2A.Z.1K101me2 peak calling process

Figure 1. ChIP analysis.

**(a)** SMYD3, H2A.ZK101me2, H2A.Z, and H2A.Zac occupancy around TSSs. SMYD3

and H2A.ZK101me2 ChIP signals were normalized to input sequences. H2A.Z and

H2A.Zac datasets were obtained from the public database (GEO accession number GSM1059388). **(b and c)** The workflow of ChIP-seq analysis. Details were described in the results, and the materials and methods Excel files were kept in the Lab computer.

24
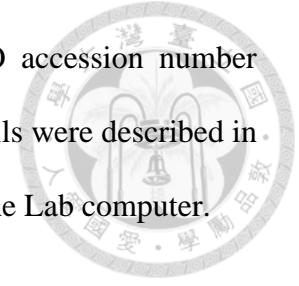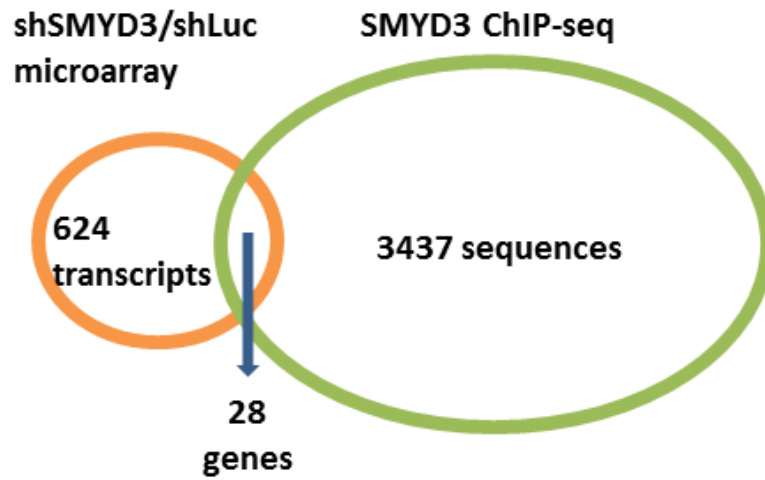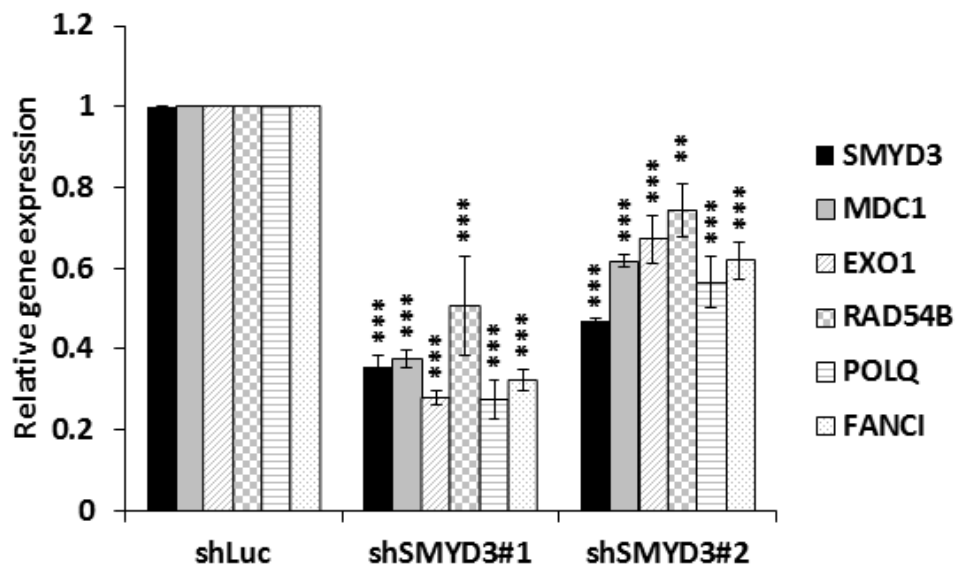
Figure 2. SMYD3 may be involved in DNA repair.

(a) 28 genes were identified in the cross-referenced data comparison between whole-genome microarray analysis of RNAs isolated from shLuc vs. shSMYD3 MCF7 cells and SMYD3 ChIP analysis. **(b)** The confirmation of the microarray analyses for the

25

expressions of candidate genes in the shSMYD3/shLuc dataset using qRT-PCR. The fold

changes of each gene expression in the microarray data were listed in Table 1. **, $P <$

0.01, ***, $P < 0.001$ vs. shLuc control. All values in the histograms were means ± SD of

triplicates and data were representative of n ≥ 3 for each experiment.

26

a

shLuc/shSMYD3

(-log *p*-value)

1. cell cycle
2. DNA metabolic process
3. response to DNA damage stimulus
4. cell proliferation
5. macromolecular complex subunit organization
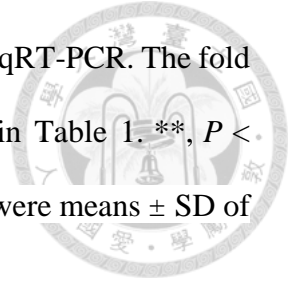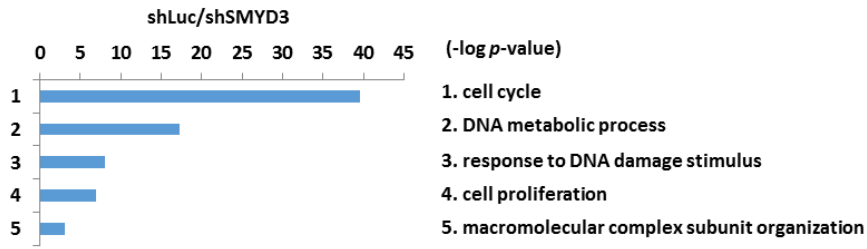
b

c

Figure 3. SMYD3 is required for DNA repair machinery.

(a) Whole-genome microarray analysis of RNAs isolated from shLuc vs. shSMYD3 MCF7 cells was conducted. With a cut-off of absolute normalized fold change $\leq$ 0.5 (log2 normalized ratios < -1), the list of down-regulated genes was further categorized by the DAVID v6.8 Gene Ontology program to reveal their gene ontology process. **(b)** γH2A.X foci formation at indicated times after 1.67 Gy IR treatment in shLuc or shSMYD3 MCF7 cells. *, $P < 0.05$. **, $P < 0.01$. **(c)** γH2A.X foci formation at indicated times after 1.67 Gy IR treatment in shLuc or shSMYD3 MCF7 cells complemented with the vector control, SMYD3 or mutant SMYD3[Y239F]. *, $P < 0.05$. **, $P < 0.01$. ***, $P < 0.001$.

28
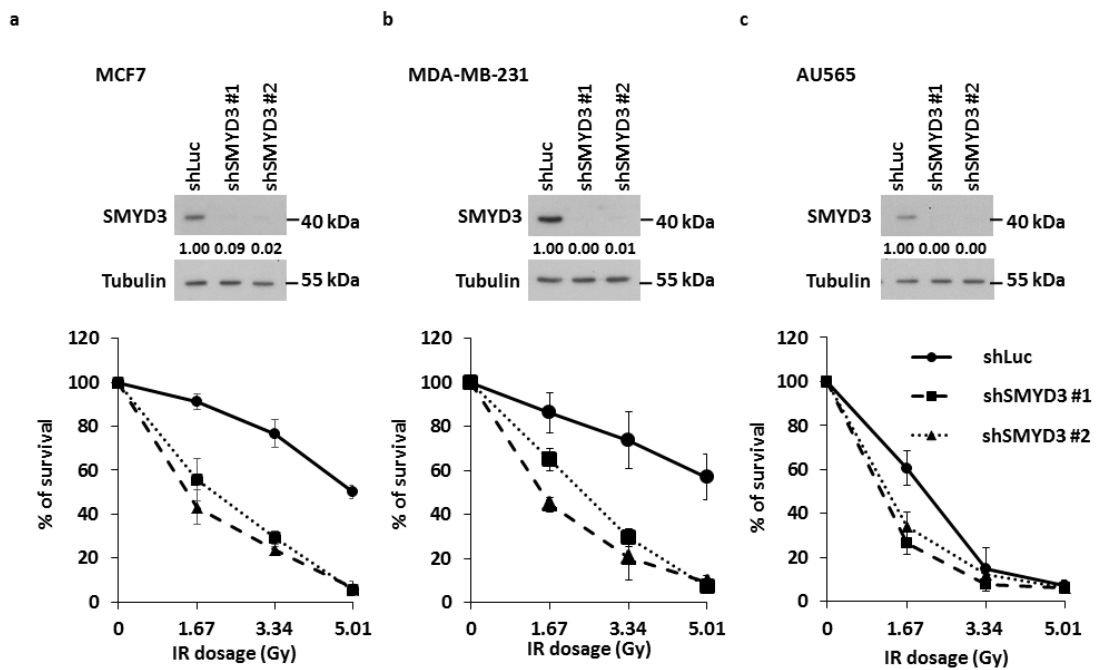
Figure 4. SMYD3-depleted cells are hypersensitive to IR stress.

(**a-c**) Clonogenic survivals of shLuc and shSMYD3 MCF7 (**a**), MDA-MB-231 (**b**) and

AU565 cells (**c**) treated with indicated dosages of IR. All values in the diagrams were

means ± SD of triplicates and data were representative of n ≥ 3 for each experiment.

29

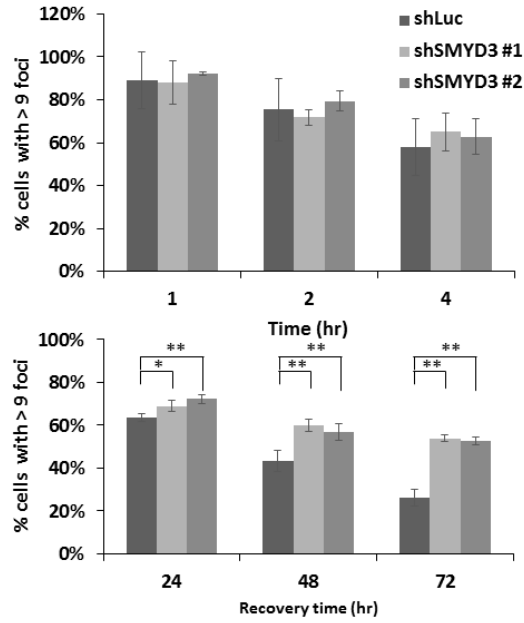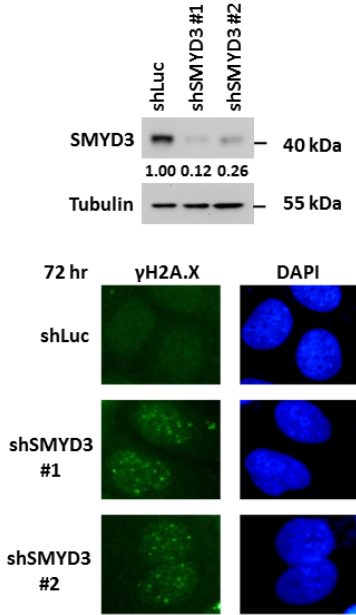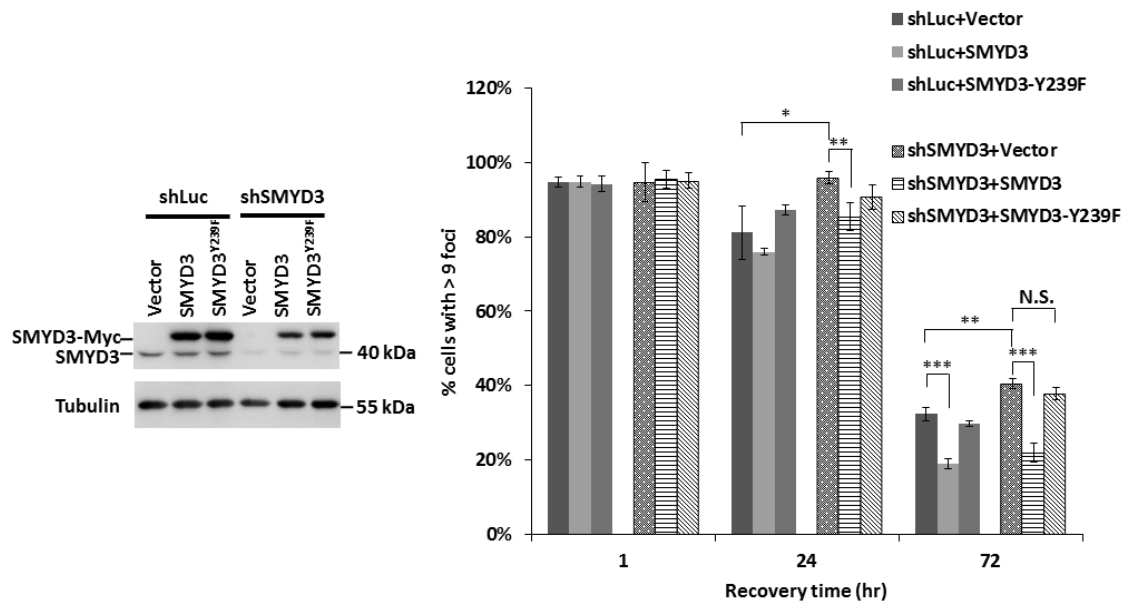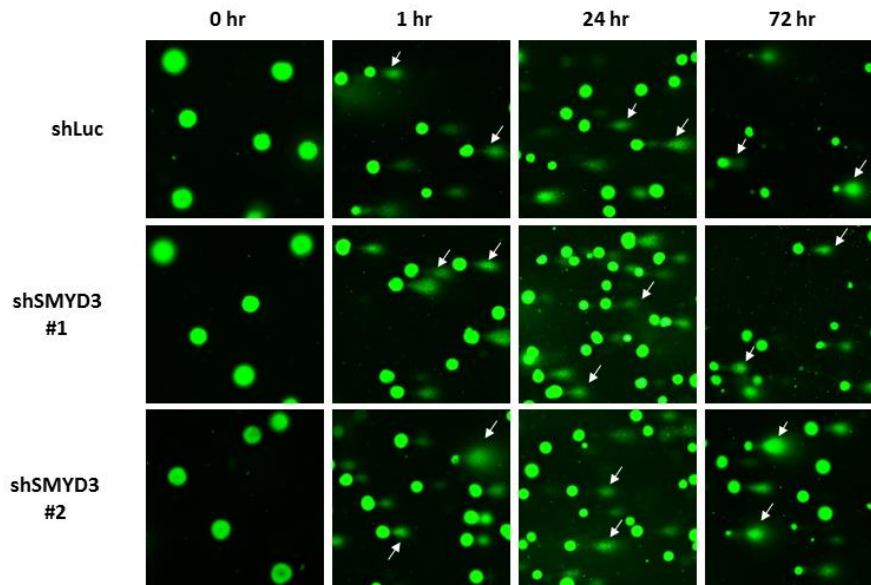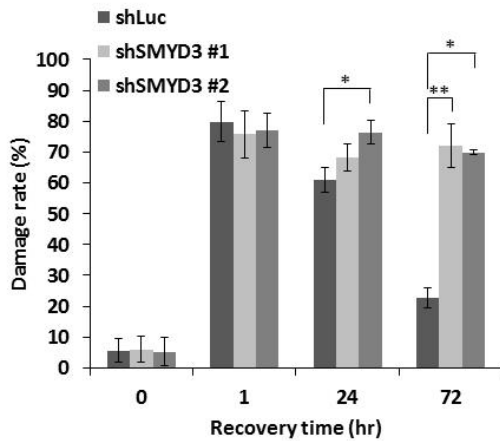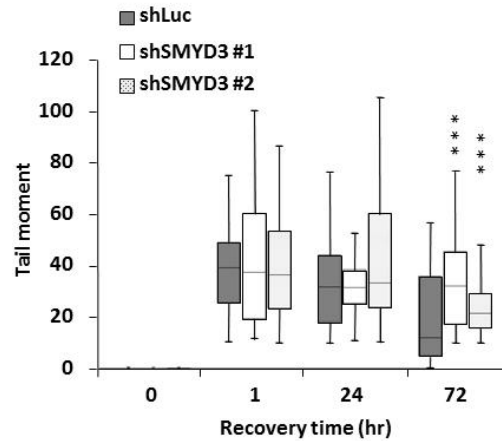**Figure 5. SMYD3 deficiency increases the ratio of DNA damage after IR treatment.**

shLuc and shSMYD3 MCF7 cells were treated with 1.67 Gy IR and then analyzed for DNA damage by comet assay at indicated times. **(a)** Representative images with white arrows point to comet tails, which signify DNA damage. **(b)** Quantification of the percentage of nuclei with comet tails. **(c)** Quantification of comet tail moment using the

30

CometScore software. Data are presented as a quantile box plot. The middle lines in boxes indicate the median; upper and lower box edges indicate the 25th and 75th percentiles; and bars indicate the 10th and 90th percentiles. Statistical analysis was performed by one-way analysis variance (35). ***, $P < 0.001$.
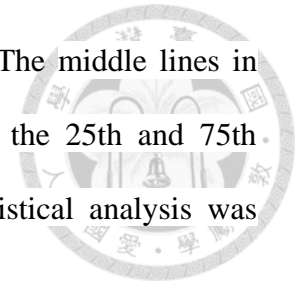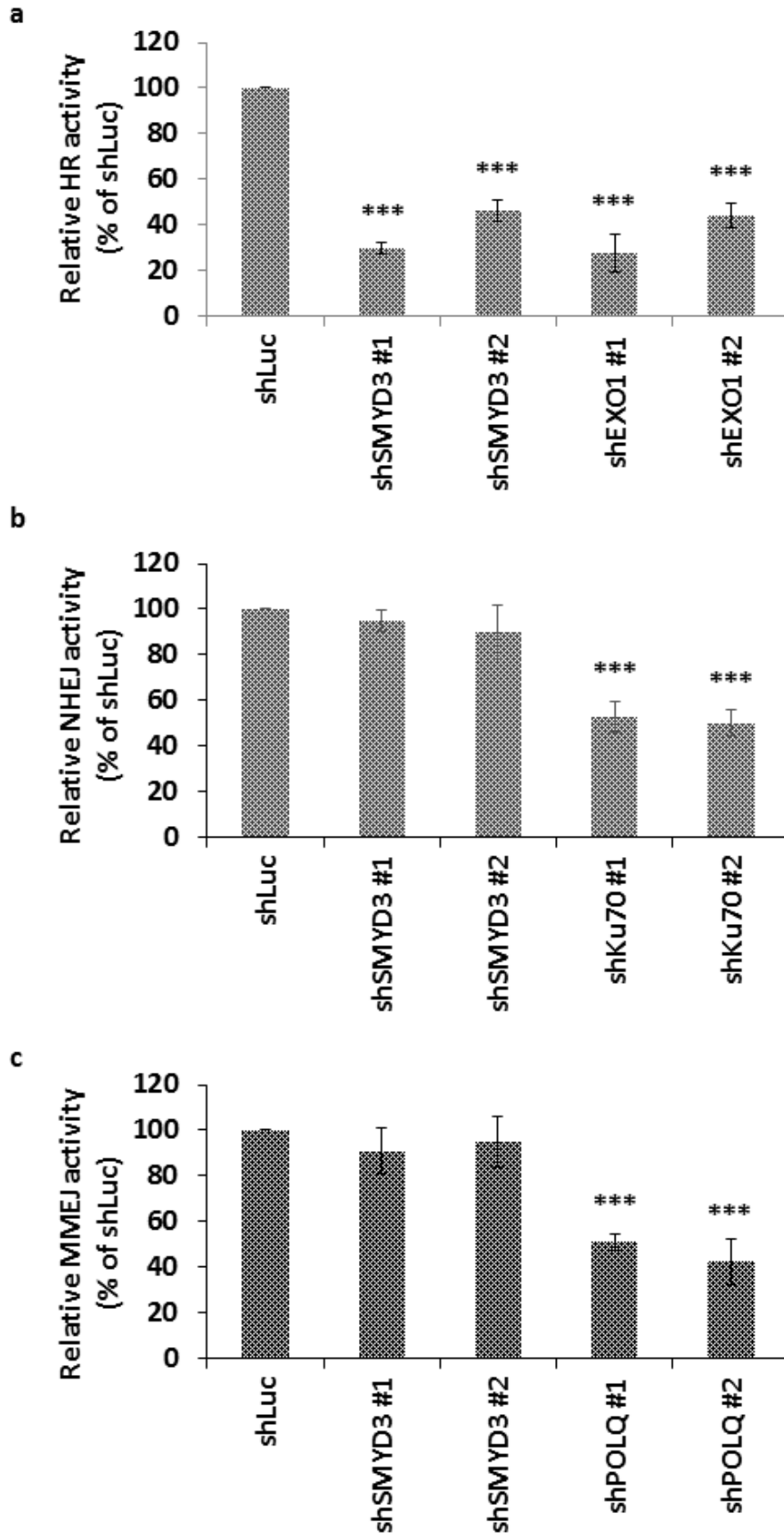
Figure 6. SMYD3 mediates the HR pathway.

(a) Relative HR activity of shLuc, shSMYD3, and shEXO1 MCF7 cells, which was indicated by the percentage of GFP-positive cells relative to the shLuc control. The EXO1 knockdown cells were used as a positive control. ***, $P < 0.001$ vs. shLuc control. **(b)** Relative NHEJ activity of shLuc, shSMYD3 and shKu70 cells. NHEJ activity was measured by normalizing the luciferase activity to the renilla activity. The Ku70-knockdown cells were used as a positive control. ***, $P < 0.001$ vs. shLuc control.

(c) Relative MMEJ activity of shLuc, shSMYD3 and shPOLQ MCF7 cells, which was indicated by the percentage of GFP-positive cells relative to the shLuc control. The POLQ-knockdown cells were used as a positive control. ***, $P < 0.001$ vs. shLuc control.

33
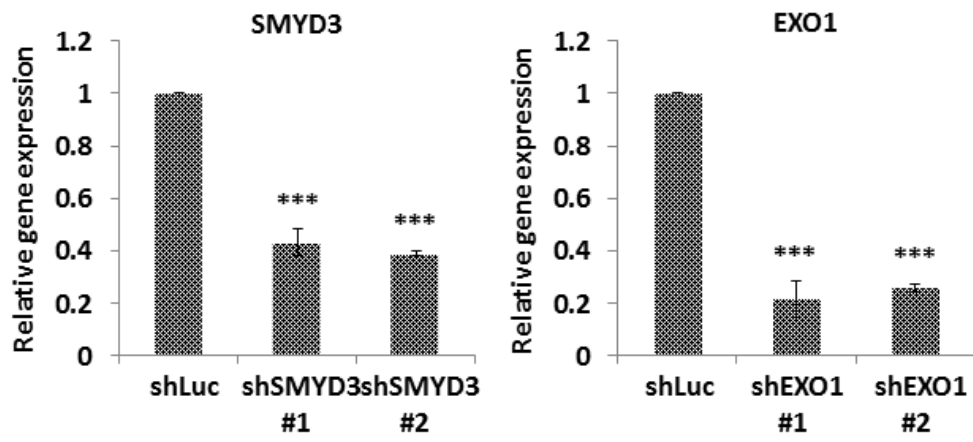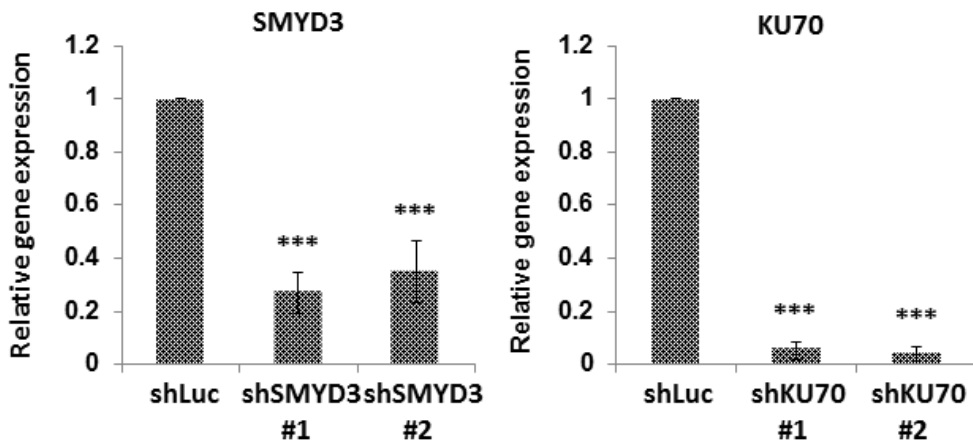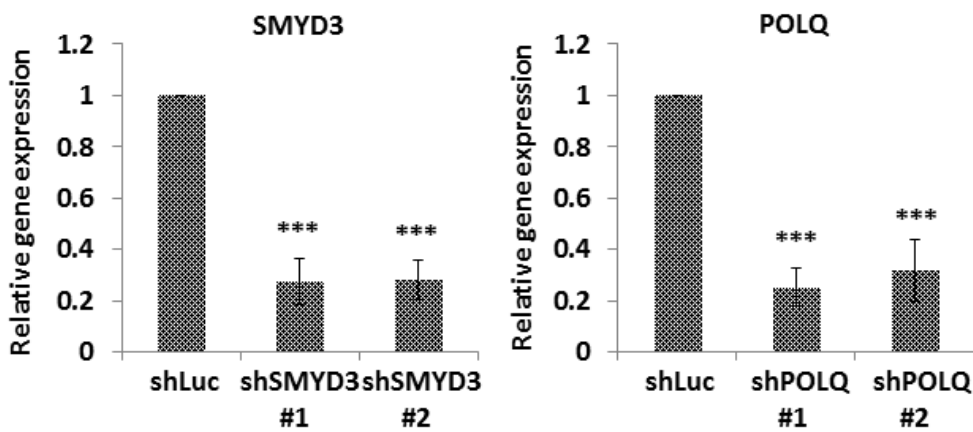
Figure 7. The knockdown efficiency of each knockdown clones used.

The knockdown efficiency was validated by qRT-PCR six days after selected with puromycin. **(a)** Knockdown efficiency of *SMYD3* and *EXO1* in MCF7/DR-GFP cells. **(b)** Knockdown efficiency of *SMYD3* and *Ku70* in MCF7 cells. **(c)** Knockdown efficiency of *SMYD3* and *POLQ* in MCF7 cells. ***, $P < 0.001$ vs. shLuc control. All values in the histograms *were* means ± SD of *triplicates* and data were representative of n ≥ 3 for each experiment.

35

a
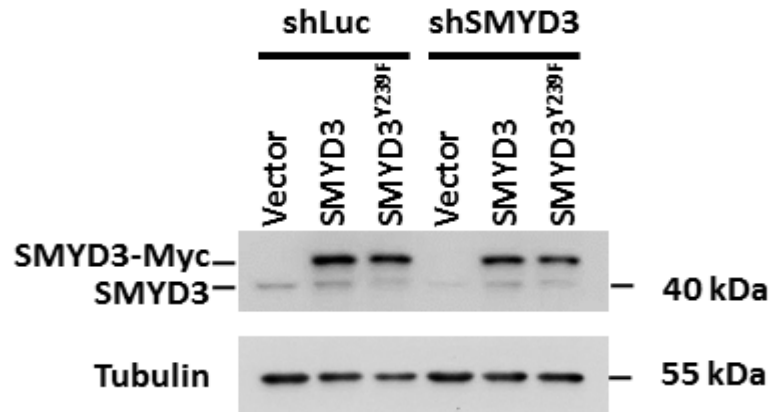
b
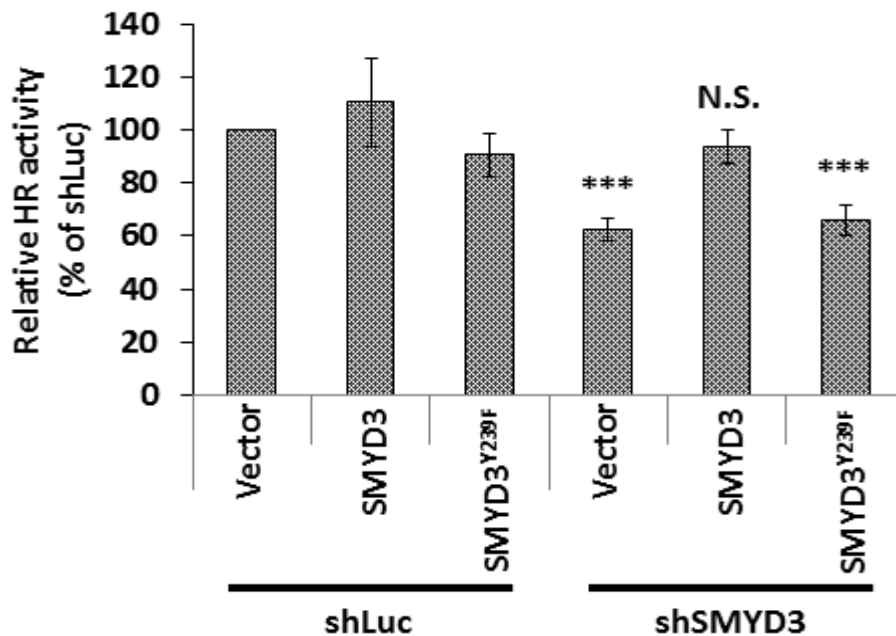
Figure 8. Complementation of SMYD3 in shSMYD3 cells is able to recover its HR deficiency.

(a) Western blotting of cells subjected to Luc or SMYD3 knockdown and complemented with vector control or Myc-tagged SMYD3 or mutant SMYD3^{Y239F}. **(b)** Exogenous expression of SMYD3, but not the mutant SMYD3^{Y239F}, restored the HR activity. ***, $P < 0.01$ vs. shLuc with the expression of the vector plasmid control. N.S.:

36

not significant.

37

Fold changes in array data	SMYD3	TRIP13	MDC1	BARD1	UHRF1	FANCD2	BRCA2	EXO1	RAD18	RAD51	CHEK1	LIG1	RAD51AP1	RAD54B
	0.33	0.33	0.35	0.38	0.41	0.47	0.47	0.47	0.47	0.47	0.50	0.50	0.50	0.50
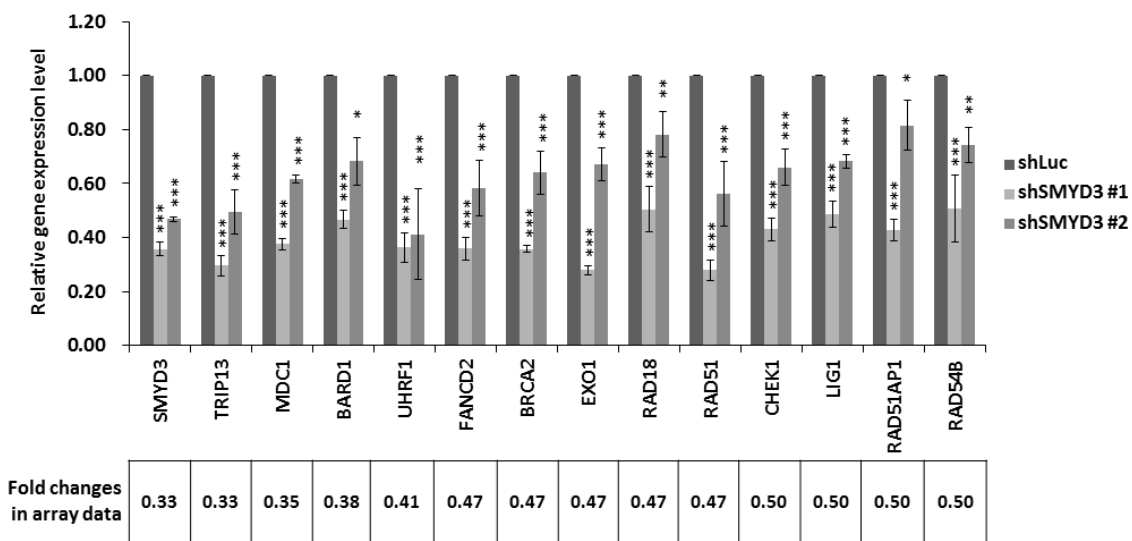
Figure 9. SMYD3 knockdown downregulates HR gene expressions.

The confirmation of the microarray analyses for the expressions of candidate genes in the shSMYD3/shLuc dataset using qRT-PCR. The fold changes of each gene expression in the microarray data were listed below. *, $P < 0.05$, **, $P < 0.01$, ***, $P < 0.001$ vs. shLuc control. All values in the histograms were means ± SD of triplicates and data were representative of n ≥ 3 for each experiment.

38
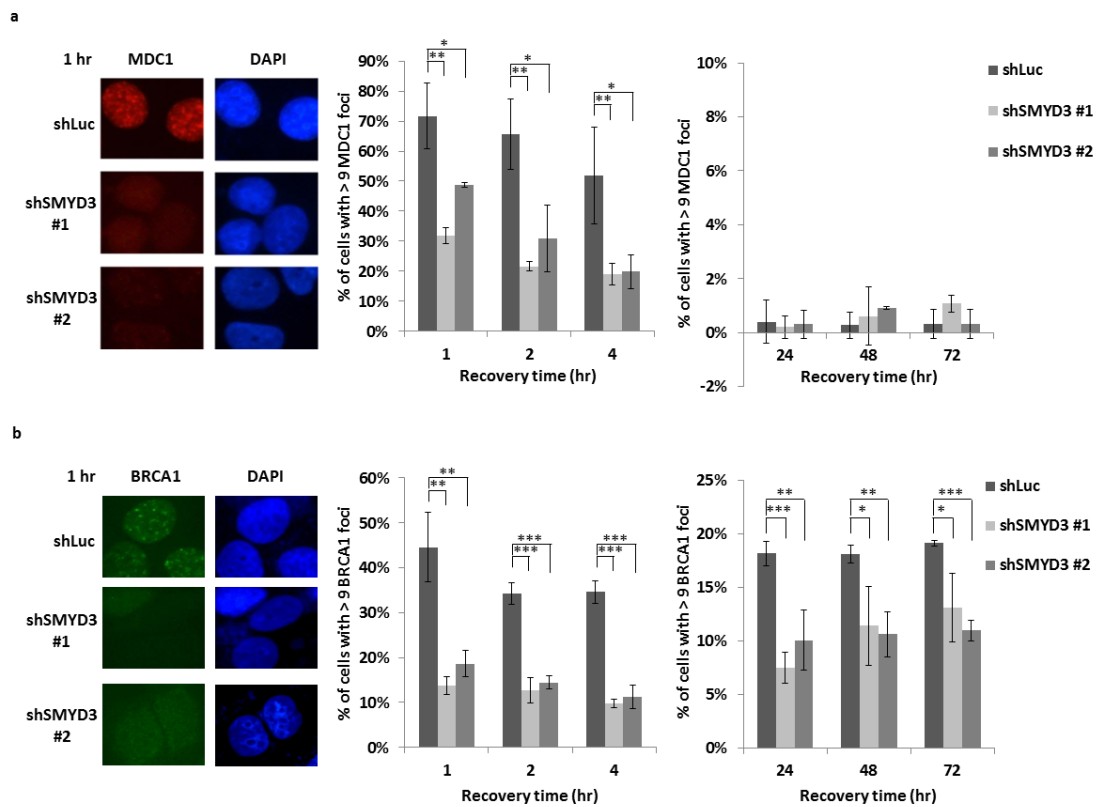
Figure 10. SMYD3 knockdown downregulates the expression of DNA repair foci.

(a and b) DNA repair foci formation of MDC1 foci **(a)** and BRCA1 foci **(b)** formation at indicated times after 1.67 Gy IR treatment in shLuc or shSMYD3 MCF7 cells. *, $P <$ 0.05. **, $P < 0.01$. ***, $P < 0.001$. All values in the histograms were means ± SD of triplicates and data were representative of n ≥ 3 for each experiment.
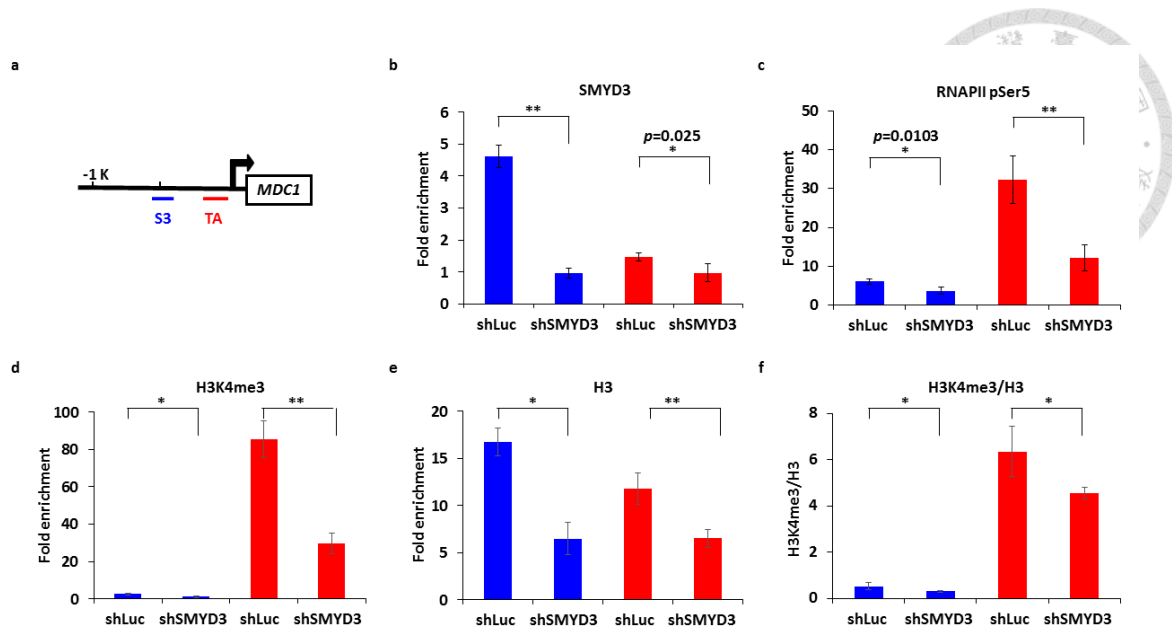
Figure 11. SMYD3 regulates the expression of *MDC1* through methylating histone H3K4.

(a) ChIP assay was performed in MCF7 cells using specific antibodies. The examined positions at the *MDC1* locus were indicated, in which region S3 and region TA for the predicted SMYD3 and TATA box binding sites, respectively. **(b-e)** ChIP assays were performed with SMYD3-repressed MCF7 cells using specific antibodies shown at the top of the histogram. Fold enrichment of each antibody compared with IgG was shown. **(f)** Ratios of H3K4me3/H3 ChIP signals were displayed. In **(b-f)**, immunoprecipitated chromatin was quantified by qRT-PCR. *, $P < 0.05$. **, $P < 0.01$. All values in the histograms were means ± SD of triplicates and data were representative of n ≥ 3 for each experiment.

40
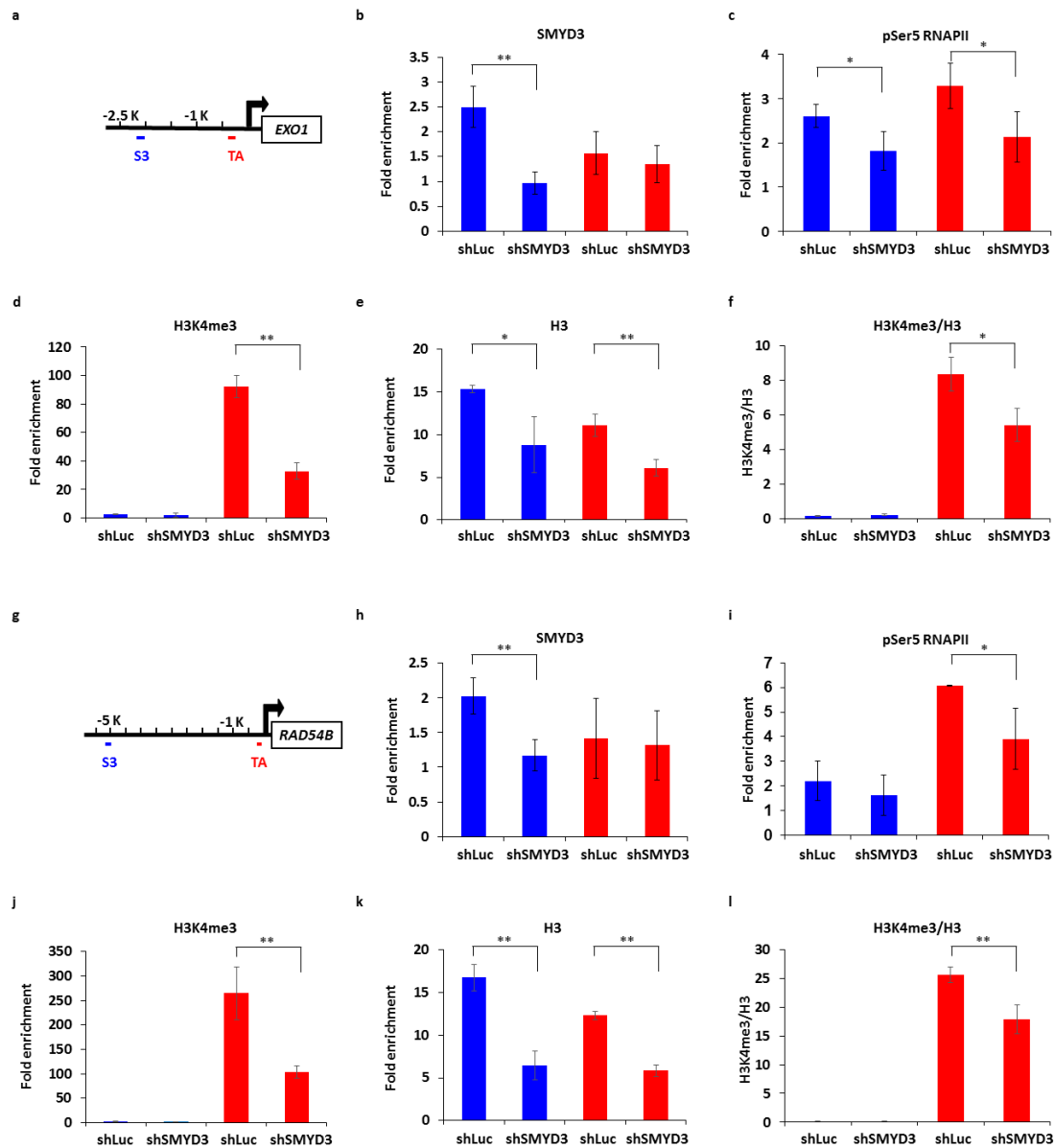
**Figure 12. SMYD3 activates the expression of *EXO1* and *RAD54B* through methylating histone H3K4.**

(a and g) ChIP assays were performed in MCF7 cells using specific antibodies. The examined positions at *EXO1* and *RAD54B* loci were indicated, in which region S3 and region TA are predicted SMYD3 and TATA box binding sites, respectively. **(b-e and h-k)** ChIP assays were performed with SMYD3-repressed MCF7 cells using specific

41

antibodies indicated at the top of the histogram. Fold enrichment of each antibody compared with IgG was shown. **(f and l)** Ratios of H3K4me3/H3 ChIP signals were shown. In **b-f** and **h-l**, immunoprecipitated chromatin was quantified by qRT-PCR. *, $P < 0.05$. **, $P < 0.01$. All values in the histograms were means ± SD of triplicates and data were representative of n ≥ 3 for each experiment.
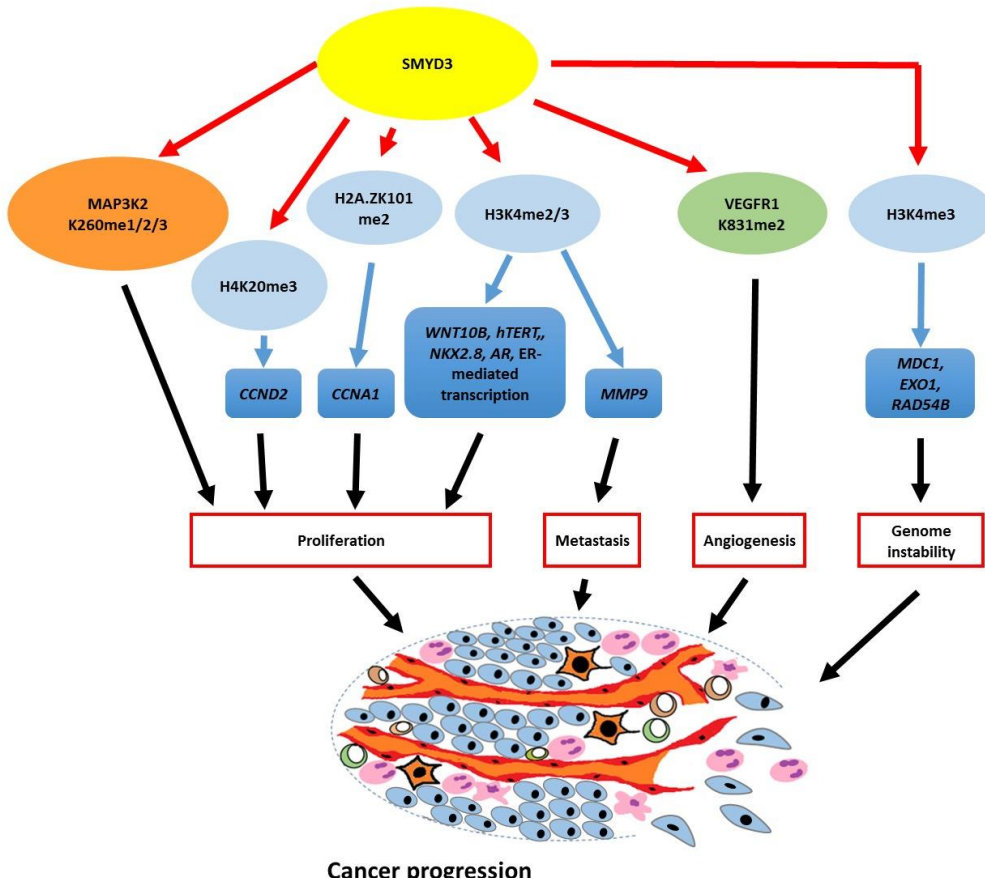
**Figure 13. SMYD3 methylates histone and non-histone substrates to regulate different pathways that are important for hallmarks of cancer.**

SMYD3 directly methylates proteins that are involved in cell proliferation and angiogenesis. Moreover, SMYD3 methylates histones that are widely spread on chromatin to facilitate transcription of target genes. Hence, the high level of SMYD3 stimulates cancer development.

Chapter 6. Table.

Table 1. Down-regulated genes in shSMYD3/shLuc array data (genes with < 0.5 fold differences) cross-referenced with SMYD3 ChIP-seq data.

Log ratio	Ratio	Gene Title/Gene Symbol
-1.0	0.50	Anillin, Actin Binding Protein/*ANLN*
-1.0	0.50	DNA (Cytosine-5-)-Methyltransferase 3 Beta/*DNMT3B*
-1.0	0.50	Geminin, DNA replication inhibitor/*GMNN*
-1.0	0.50	5'-nucleotidase, cytosolic II/*NT5C2*
-1.0	0.50	POC1 centriolar protein A/*POC1A*
-1.0	0.50	RAD54 homolog B (S. cerevisiae)/*RAD54B*
-1.0	0.50	T-cell lymphoma invasion and metastasis 1/*TIAM1*
-1.1	0.47	Exonuclease 1/*EXO1*
-1.1	0.47	Fanconi anemia, complementation group I/*FANCI*
-1.1	0.47	IQ motif containing GTPase activating protein 3/*IQGAP3*
-1.1	0.47	Opa interacting protein 5/*OIP5*
-1.1	0.47	Polymerase (DNA directed), alpha 2, accessory subunit/*POLA2*
-1.1	0.47	Polymerase (DNA directed), theta/*POLQ*
-1.2	0.44	Low density lipoprotein receptor-related protein 1/*LRP1*
-1.2	0.44	Protein kinase, cAMP-dependent, regulatory, type II, beta/*PRKAR2B*
-1.2	0.44	Ret proto-oncogene/*RET*
-1.2	0.44	Solute carrier family 7 (amino acid transporter light chain, L system), member 5/*SLC7A5*
-1.3	0.41	ST6 (alpha-N-acetyl-neuraminyl-2,3-beta-galactosyl-1,3)-N-acetylgalactosaminide alpha-2,6-sialyltransferase 2/

		ST6GALNAC2
-1.4	0.38	Integrin, alpha 9/*ITGA9*
-1.4	0.38	V-myb avian myeloblastosis viral oncogene homolog-like 2/*MYBL2*
-1.5	0.35	Glutathione peroxidase 2 (gastrointestinal)/*GPX2*
-1.5	0.35	Mediator of DNA-damage checkpoint 1/*MDC1*
-1.5	0.35	Nuclear receptor subfamily 5, group A, member 2/*NR5A2*
-1.6	0.33	Cytochrome B Reductase 1/*CYBRD1*
-1.6	0.33	UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase 12 (GalNAc-T12)/*GALNT12*
-1.7	0.31	Glutamate decarboxylase 2 (pancreatic islets and brain, 65kDa)/*GAD2*
-1.8	0.29	EGF containing fibulin-like extracellular matrix protein 1/*EFEMP1*
-1.8	0.29	Serpin peptidase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 1/*SERPINA1*

45
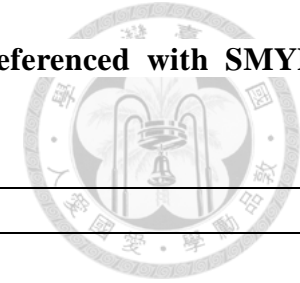
Table 2. Up-regulated genes in shSMYD3/shLuc array data (genes with < 0.5 fold differences) cross-referenced with SMYD3 ChIP-seq data.

Log ratio	Ratio	Gene Title/Gene Symbol
3.6	12.13	DnaJ (Hsp40) homolog, subfamily B, member 8/*DNAJB8*
2.8	6.96	Glycerophosphodiester phosphodiesterase domain containing 1/*GDPD1*
2.7	6.50	Cytoplasmic polyadenylation element binding protein 1/*CPEB1*
2.6	6.06	Sema domain, transmembrane domain (TM), and cytoplasmic domain, (semaphorin) 6A/*SEMA6A*
2.5	5.66	Hydroxysteroid (17-beta) dehydrogenase 14/*HSD17B14*
1.9	3.73	Tumor necrosis factor (ligand) superfamily, member 10/*TNFSF10*
1.8	3.48	RAR-related orphan receptor C/*RORC*
1.5	2.83	CD36 molecule (thrombospondin receptor)/*CD36*
1.4	2.64	HCLS1 binding protein 3/*HS1BP3*
1.4	2.64	Ring finger protein 43/*RNF43*
1.3	2.46	Osteopetrosis associated transmembrane protein 1/*OSTM1*
1.2	2.30	Protein kinase (cAMP-dependent, catalytic) inhibitor gamma/*PKIG*
1.2	2.30	RAB7, member RAS oncogene family-like 1/*RAB7L1*
1.2	2.30	Yippee-like 3 (Drosophila)/*YPEL3*
1.2	2.30	Zinc finger, DHHC-type containing 11/*ZDHHC11*
1.1	2.14	BCL2/adenovirus E1B 19kD interacting protein like/*BNIPL*
1.1	2.14	Cytochrome c oxidase subunit VIIb/*COX7B*
1.1	2.14	DENN/MADD domain containing 1B/*DENND1B*
1.1	2.14	Fibrosin-like 1/*FBRSL1*
1.1	2.14	Radial spoke head 1 homolog (Chlamydomonas)/*RSPH1*

46

1.1	2.14	TM2 domain containing 1/*TM2D1*
1.1	2.14	Zinc finger protein 44/*ZNF44*
1.0	2.00	Aminoadipate-semialdehyde synthase/*AASS*
1.0	2.00	B-cell CLL/lymphoma 11A (zinc finger protein)/*BCL11A*
1.0	2.00	Diacylglycerol O-acyltransferase 1/*DGAT1*
1.0	2.00	ORAI calcium release-activated calcium modulator 3/*ORAI3*
1.0	2.00	Peroxisomal biogenesis factor 19/*PEX19*
1.0	2.00	Retinol saturase (all-trans-retinol 13,14-reductase)/*RETSAT*
1.0	2.00	SUMO1/sentrin/SMT3 specific peptidase 2/*SENP2*
1.0	2.00	Sbiquitin-conjugating enzyme E2I/*UBE2I*

Table 3. Down-regulated genes in H2A.Z.1^{WT}/H2A.Z.1^{K101Q} array data (genes with < 0.5 fold differences) cross-referenced with H2A.Z.1K101me2 ChIP-seq data.
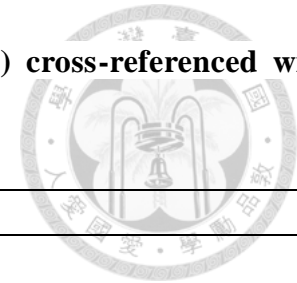
Log ratio	Ratio	Gene Title/Gene Symbol
-1.1	0.47	Cadherin 18, type 2/*CDH18*
-1.3	0.41	Heat shock 22kDa protein 8/*HSPB8*
-2.3	0.20	DnaJ (Hsp40) homolog, subfamily C, member 2/*DNAJC2*
-2.3	0.20	Hemoglobin, alpha 1/*HBA1*
-2.4	0.19	ELL associated factor 2/*EAF2*
-3.3	0.10	Apolipoprotein B mRNA editing enzyme, catalytic polypeptide 1/*APOBEC1*
-3.4	0.09	Sperm associated antigen 17/*SPAG17*
-3.6	0.08	Lysophosphatidic acid receptor 5/*LPAR5*

48

Table 4. Down-regulated genes in H2A.Z.1^{WT}/H2A.Z.1^{K101Q} array data (genes with < 0.5 fold differences) cross-referenced with H2A.Z.1K101me2 ChIP-seq data.
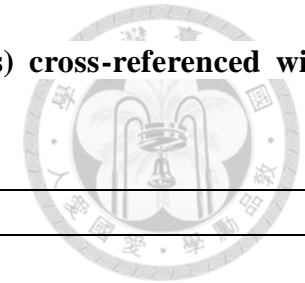
Log ratio	Ratio	Gene Title/Gene Symbol
4.3	19.70	Reversion-inducing-cysteine-rich protein with kazal motifs/*RECK*
3.5	11.31	Zyg-11 related, cell cycle regulator/*ZER1*
3.0	8.00	Developmental pluripotency associated 2/*DPPA2*
2.9	7.46	SRSF protein kinase 2/*SRPK2*
2.3	4.92	Reticulon 1/*RTN1*
2.2	4.59	B-cell CLL/lymphoma 9/*BCL9*
1.4	2.64	Ankyrin repeat domain 45/*ANKRD45*
1.4	2.64	Proline-rich coiled-coil 1/*PRRC1*
1.3	2.46	Alanyl (membrane) aminopeptidase/*ANPEP*
1.3	2.46	Tripartite motif containing 16-like/*TRIM16L*
1.3	2.46	Xenotropic and polytropic retrovirus receptor 1/*XPR1*
1.2	2.30	ARFGEF family member 3/*KIAA1244* (*ARFGEF3*)
1.1	2.14	Karyopherin alpha 5 (importin alpha 6)/*KPNA5*
1.0	2.00	Carbonic anhydrase VA, mitochondrial/*CA5A*
1.0	2.00	Early growth response 1/*EGR1*

Table 5. GO analysis of down-regulated genes in shSMYD3/shLuc array data (genes with < 0.5 fold differences) cross-referenced with SMYD3 ChIP-seq data.

Term	Genes	Count	*p*-value	Benjamini
DNA metabolic process	*DNMT3B, FANCI, RAD54B, EXO1, MDC1, POLA2, POLQ*	7	2.0E-4	7.6E-2
DNA repair	*MDC1, EXO1, RAD54B, POLQ, FANC,*	5	1.5E-3	2.5E-1
Cell cycle	*FANCI, OIP5, RAD54B, ANLN, EXO1, GMNN, MDC1*	7	1.9E-3	2.2E-1
M phase	*OIP5, RAD54B, ANLN, EXO1*	4	2.0E-2	7.3E-1

50
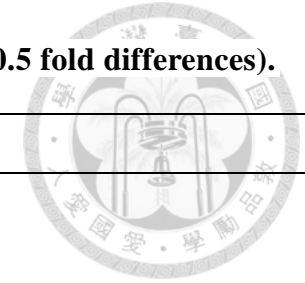
Table 6. Down-regulated DNA damage stimulus response genes in shSMYD3/shLuc array data (genes with < 0.5 fold differences).

Function	Gene Title/Gene Symbol
Antioxidant	Thioredoxin domain containing 12/*TXNDC12*
Base excision repair	Nei like DNA glycosylase 3/*NEIL3*
Base excision repair	Nudix hydrolase 1/*NUDT1*
Base excision repair	DNA polymerase delta 1, catalytic subunit/*POLD1*
Cell cycle	Cyclin A2/*CCNA2*
Cell cycle	Cyclin dependent kinase 1/*CDK1*
Checkpoint effector	G2 and S-phase expressed 1/*GTSE1*
DNA replication	Thymidylate synthetase/*TYMS*
DNA replication	Minichromosome maintenance complex component 7/*MCM7*
DNA replication	DNA polymerase epsilon 2, accessory subunit/*POLE2*
DNA replication and repair	Topoisomerase (DNA) II alpha/*TOP2A*
G2/M checkpoint	Denticleless E3 ubiquitin protein ligase homolog/*DTL*
HR	BRCA1 associated RING domain 1/*BARD1*
HR	RAD54 homolog B (S. cerevisiae)/*RAD54B*
HR	RAD18, E3 ubiquitin protein ligase/*RAD18*
HR	Ubiquitin like with PHD And Ring finger domains 1/*UHRF1*
HR	RAD51 recombinase/*RAD51*

51

HR	BRCA2, DNA repair associated/*BRCA2*
HR	Checkpoint kinase 1/*CHEK1*
HR	Exonuclease 1/*EXO1*
HR	RAD51 associated protein 1/*RAD51AP1*
HR	DNA ligase 1/*LIG1*
HR, NHEJ	Mediator Of DNA damage checkpoint 1/*MDC1*
HR, mitotic check point	Thyroid hormone receptor interactor 13/*TRIP13*
HR, interstrand cross-links	Fanconi anemia complementation group D2/*FANCD2*
Interstrand cross-links	DNA cross-link repair 1B/*DCLRE1B*
Interstrand cross-links	Fanconi anemia complementation group I/*FANCI*
Interstrand cross-links	Fanconi anemia complementation group B/*FANCB*
MMEJ	DNA polymerase theta/*POLQ*
NHEJ	Pituitary tumor-transforming 1/*PTTG1*
Postreplication repair	Ubiquitin conjugating enzyme E2 V2/*UBE2V2*
Sister chromatid cohesion	Establishment of sister chromatid cohesion N-acetyltransferase 2/*ESCO2*
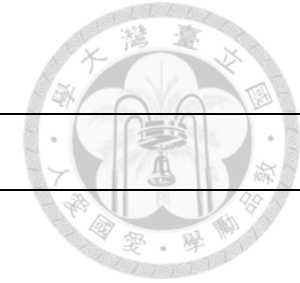
Table 7. Down-regulated HR genes in shSMYD3/shLuc array data (genes with < 0.5 fold differences).

Log ratio	Ratio	Gene Title/Gene Symbol
-1.6	0.33	SET and MYND domain containing 3/*SMYD3*
-1.6	0.33	Thyroid hormone receptor interactor 13/*TRIP13*
-1.5	0.35	Mediator of DNA-damage checkpoint 1/*MDC1*
-1.4	0.38	BRCA1 associated RING domain 1/*BARD1*
-1.3	0.41	Ubiquitin-like with PHD and ring finger domains 1/*UHRF1*
-1.2	0.44	BRCA1 associated RING domain 1/*BARD1*
-1.1	0.47	Fanconi anemia, complementation group D2/*FANCD2*
-1.1	0.47	Breast cancer 2, early onset/*BRCA2*
-1.1	0.47	Exonuclease 1/*EXO1*
-1.1	0.47	RAD18 homolog (*S. cerevisiae*)/*RAD18*
-1.1	0.47	RAD51 homolog (RecA homolog, *E. coli*) (*S. cerevisiae*)/*RAD51*
-1.0	0.50	CHK1 checkpoint homolog (*S. pombe*)/*CHEK1*
-1.0	0.50	Fanconi anemia, complementation group D2/*FANCD2*
-1.0	0.50	Ligase I, DNA, ATP-dependent/*LIG1*
-1.0	0.50	RAD51 associated protein 1/*RAD51AP1*
-1.0	0.50	RAD54 homolog B (*S. cerevisiae*)/*RAD54B*

53
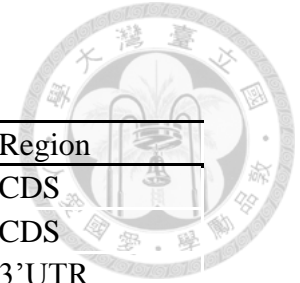
Table 8. Oligo sequences for shRNA-mediated gene knockdown

Clone ID	Gene Symbol	Target Sequence	Region
TRCN0000123290	*SMYD3*	GCTTCCCGATATCAACATCTA	CDS
TRCN0000123291	*SMYD3*	CAACTCTTTCACCATCTGTAA	CDS
TRCN0000331178	*EXO1*	TGCAGACTGCTGCAAAGCTTT	3'UTR
TRCN0000010332	*EXO1*	AATGCAGACTGCTGCAAAGCT	CDS
TRCN0000332901	*XRCC6 (KU70)*	GATGAGTCATAAGAGGATCAT	CDS
TRCN0000332902	*XRCC6 (KU70)*	CCCAAGGTTGAAGCAATGAAT	CDS
TRCN0000290546	*POLQ*	CCTTCAATCTTGCTTGCGAAA	CDS
TRCN0000290546	*POLQ*	GCTGACCAAGATTTGCTATAT	CDS

54
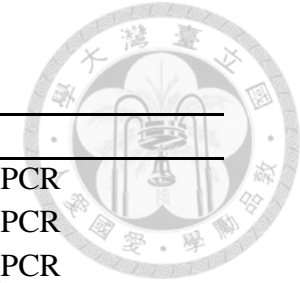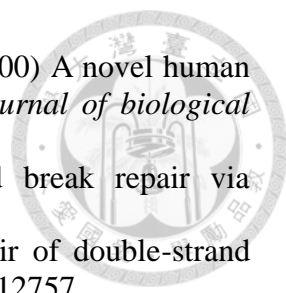
Table 9. Primers used in this study

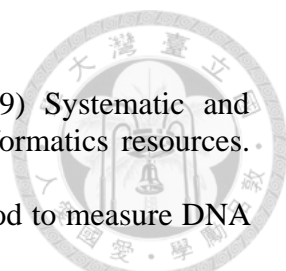| Genes	Forward sequence	Reverse sequence	Assay
SMYD3	TTACTGCGAGCAGTCCGAGACA	TTGTCCTGGGTTTGGCAACGGA	SYBR green qPCR
TRIP13	CAGCAGCACTGCAAAGAAAG	AAATCGATGGGCTGTGAGTC	SYBR green qPCR
MDC1	GCAAGATGCCACCTGCTGAGAA	GCTTCAGGTACTGTAGGAGGCA	SYBR green qPCR
BARD1	TGTGGTTTAGCCCTCGAAGT	GCCCTCTCAGAAACATCTGC	SYBR green qPCR
UHRF1	TGTGGACCATGGGAATTTTT	GGGAGCAAAGCAGTTGAGAG	SYBR green qPCR
FANCD2	TTCCAGGATGCCTTCGTAGTGG	GCAGGAGGTTTATGGCAATCCC	SYBR green qPCR
BRCA2	GGCTTCAAAAAGCACTCCAGATG	GGATTCTGTATCTCT TGACG TTCC	SYBR green qPCR
EXO1	TCGGATCTCCTAGCTTTTGGCTG	AGCTGTCTGC ACATT CCTAG CC	SYBR green qPCR
RAD18	GGATTGTCCTGTTTGCGGGGTT	GTTTTGGGCA GCGGC TTCCT TT	SYBR green qPCR
RAD51	TCTCTGGCAGTGATGTCCTGGA	TAAAGGGCGG TGGCA CTGTC TA	SYBR green qPCR
CHEK1	GTGTCAGAGTCTCCCAGTGGAT	GTTCTGGCTG AGAAC TGGAG TAC	SYBR green qPCR
LIG1	TCACAGAGGCTGAAGTGGC	TCAGGCTCTG AAACG CTTTC CG	SYBR green qPCR
RAD51AP1	CTTCTGGAAGGCAGTGATGGTG	AGAGAAGTCTTCGTCATTATCCTC	SYBR green qPCR
RAD54B	TCATGATCTG CTTGA CTGTG AG	TTTTTCCAACGAATCACCTGT	SYBR green qPCR
KU70	TGCCACAGGA AGAAG AGTTG	CTCTG GAGTT GCCAT GATTT	SYBR green qPCR
POLQ	CTTGTGGCAT CTCCT TGGAG CA	AATCC CTTGG CTGGT CTCCA TC	SYBR green qPCR
MDC1	CCTCTCAAAGTGGTGGGATT	AATTGCTTGAACCCAGAAGG	Region S3: -530~-377 bp
MDC1	AGGAGAATCGCTTGAACCTG	CTTAAAGGCTGTCCCCACCT	Region TA: -234~-49 bp
EXO1	TCACCTGAGGTTGGGAGTTC	ACTGCAACCTCTGCCTCCT	Region S3: -2278~-2114 bp
EXO1	AAGGCCCATTTTCAAGGTCT	ATTCAGTTCACGCTGGGTTC	Region TA: -386~-237 bp
RAD54B	AGACCTCCCCAGATGATTCC	CCCGAATAGCTGGGACTACA	Region S3: -5262~-5049 bp
RAD54B	TTCGTTTCTATATTCCCAGAACCT	ATGATTCGGTGTGTGCGATA	Region TA: -398~-247 bp

Chapter 7. References.

1. Rangasamy, D. (2010) Histone variant H2A.Z can serve as a new target for breast cancer therapy. *Current medicinal chemistry*, **17**, 3155-3161.
2. Chandramouli, B., Silvestri, V., Scarno, M., Ottini, L. and Chillemi, G. (2016) Smyd3 open & closed lock mechanism for substrate recruitment: The hinge motion of C-terminal domain inferred from mu-second molecular dynamics simulations. *Biochim Biophys Acta*, **1860**, 1466-1474.
3. Musselman, C.A., Lalonde, M.E., Cote, J. and Kutateladze, T.G. (2012) Perceiving the epigenetic landscape through histone readers. *Nat Struct Mol Biol*, **19**, 1218-1227.
4. Foreman, K.W., Brown, M., Park, F., Emtage, S., Harriss, J., Das, C., Zhu, L., Crew, A., Arnold, L., Shaaban, S. *et al.* (2011) Structural and functional profiling of the human histone methyltransferase SMYD3. *PLoS One*, **6**, e22290.
5. Liu, Y., Chen, W., Gaudet, J., Cheney, M.D., Roudaia, L., Cierpicki, T., Klet, R.C., Hartman, K., Laue, T.M., Speck, N.A. *et al.* (2007) Structural basis for recognition of SMRT/N-CoR by the MYND domain and its contribution to AML1/ETO's activity. *Cancer Cell*, **11**, 483-497.
6. Dillon, S.C., Zhang, X., Trievel, R.C. and Cheng, X. (2005) The SET-domain protein superfamily: protein lysine methyltransferases. *Genome Biol*, **6**, 227.
7. Hamamoto, R., Furukawa, Y., Morita, M., Iimura, Y., Silva, F.P., Li, M., Yagyu, R. and Nakamura, Y. (2004) SMYD3 encodes a histone methyltransferase involved in the proliferation of cancer cells. *Nat Cell Biol*, **6**, 731-740.
8. Spellmon, N., Holcomb, J., Trescott, L., Sirinupong, N. and Yang, Z. (2015) Structure and function of SET and MYND domain-containing proteins. *Int J Mol Sci*, **16**, 1406-1428.
9. Van Aller, G.S., Reynoird, N., Barbash, O., Huddleston, M., Liu, S., Zmoos, A.F., McDevitt, P., Sinnamon, R., Le, B., Mas, G. *et al.* (2012) Smyd3 regulates cancer cell phenotypes and catalyzes histone H4 lysine 5 methylation. *Epigenetics : official journal of the DNA Methylation Society*, **7**, 340-343.
10. Mazur, P.K., Reynoird, N., Khatri, P., Jansen, P.W., Wilkinson, A.W., Liu, S., Barbash, O., Van Aller, G.S., Huddleston, M., Dhanak, D. *et al.* (2014) SMYD3 links lysine methylation of MAP3K2 to Ras-driven cancer. *Nature*, **510**, 283-287.
11. Bai, H., Li, Y., Gao, H., Dong, Y., Han, P. and Yu, H. (2016) Histone methyltransferase SMYD3 regulates the expression of transcriptional factors during bovine oocyte maturation and early embryonic development. *Cytotechnology*, **68**, 849-859.
12. Hamamoto, R., Silva, F.P., Tsuge, M., Nishidate, T., Katagiri, T., Nakamura, Y. and Furukawa, Y. (2006) Enhanced SMYD3 expression is essential for the growth of breast cancer cells. *Cancer Sci*, **97**, 113-118.
13. Giakountis, A., Moulos, P., Sarris, M.E., Hatzis, P. and Talianidis, I. (2017) Smyd3-associated regulatory pathways in cancer. *Semin Cancer Biol*, **42**, 70-80.
14. Martin, C. and Zhang, Y. (2005) The diverse functions of histone lysine methylation. *Nat Rev Mol Cell Biol*, **6**, 838-849.
15. Tsai, C.H., Chen, Y.J., Yu, C.J., Tzeng, S.R., Wu, I.C., Kuo, W.H., Lin, M.C.,

Chan, N.L., Wu, K.J. and Teng, S.C. (2016) SMYD3-Mediated H2A.Z.1 Methylation Promotes Cell Cycle and Cancer Proliferation. *Cancer Res*, **76**, 6043-6053.

16. Cock-Rada, A.M., Medjkane, S., Janski, N., Yousfi, N., Perichon, M., Chaussepied, M., Chluba, J., Langsley, G. and Weitzman, J.B. (2012) SMYD3 promotes cancer invasion by epigenetic upregulation of the metalloproteinase MMP-9. *Cancer Res*, **72**, 810-820.

17. Liu, C., Fang, X., Ge, Z., Jalink, M., Kyo, S., Bjorkholm, M., Gruber, A., Sjoberg, J. and Xu, D. (2007) The telomerase reverse transcriptase (hTERT) gene is a direct target of the histone methyltransferase SMYD3. *Cancer Res*, **67**, 2626-2631.

18. Kunizaki, M., Hamamoto, R., Silva, F.P., Yamaguchi, K., Nagayasu, T., Shibuya, M., Nakamura, Y. and Furukawa, Y. (2007) The lysine 831 of vascular endothelial growth factor receptor 1 is a novel target of methylation by SMYD3. *Cancer Res*, **67**, 10759-10765.

19. Hamamoto, R., Furukawa, Y., Morita, M., Iimura, Y., Silva, F.P., Li, M., Yagyu, R. and Nakamura, Y. (2004) SMYD3 encodes a histone methyltransferase involved in the proliferation of cancer cells. *Nature cell biology*, **6**, 731-740.

20. Proserpio, V., Fittipaldi, R., Ryall, J.G., Sartorelli, V. and Caretti, G. (2013) The methyltransferase SMYD3 mediates the recruitment of transcriptional cofactors at the myostatin and c-Met genes and regulates skeletal muscle atrophy. *Genes Dev*, **27**, 1299-1312.

21. Lee, S.J. (2004) Regulation of muscle mass by myostatin. *Annu Rev Cell Dev Biol*, **20**, 61-86.

22. Hanahan, D. and Weinberg, R.A. (2011) Hallmarks of cancer: the next generation. *Cell*, **144**, 646-674.

23. Figueroa-Gonzalez, G. and Perez-Plasencia, C. (2017) Strategies for the evaluation of DNA damage and repair mechanisms in cancer. *Oncol Lett*, **13**, 3982-3988.

24. Symington, L.S. and Gautier, J. (2011) Double-strand break end resection and repair pathway choice. *Annu Rev Genet*, **45**, 247-271.

25. Sartori, A.A., Lukas, C., Coates, J., Mistrik, M., Fu, S., Bartek, J., Baer, R., Lukas, J. and Jackson, S.P. (2007) Human CtIP promotes DNA end resection. *Nature*, **450**, 509-514.

26. Tran, P.T., Erdeniz, N., Symington, L.S. and Liskay, R.M. (2004) EXO1-A multi-tasking eukaryotic nuclease. *DNA Repair (Amst)*, **3**, 1549-1559.

27. Wang, A.Y., Aristizabal, M.J., Ryan, C., Krogan, N.J. and Kobor, M.S. (2011) Key Functional Regions in the Histone Variant H2A.Z C-Terminal Docking Domain. *Mol Cell Biol*, **31**, 3871-3884.

28. Thorslund, T., McIlwraith, M.J., Compton, S.A., Lekomtsev, S., Petronczki, M., Griffith, J.D. and West, S.C. (2010) The breast cancer tumor suppressor BRCA2 promotes the specific targeting of RAD51 to single-stranded DNA. *Nat Struct Mol Biol*, **17**, 1263-1265.

29. Tanaka, K., Kagawa, W., Kinebuchi, T., Kurumizaka, H. and Miyagawa, K. (2002) Human Rad54B is a double-stranded DNA-dependent ATPase and has biochemical properties different from its structural homolog in yeast, Tid1/Rdh54. *Nucleic acids research*, **30**, 1346-1353.

30. Tanaka, K., Hiramoto, T., Fukuda, T. and Miyagawa, K. (2000) A novel human rad54 homologue, Rad54B, associates with Rad51. *The Journal of biological chemistry*, **275**, 26316-26321.

31. Davis, A.J. and Chen, D.J. (2013) DNA double strand break repair via non-homologous end-joining. *Transl Cancer Res*, **2**, 130-143.

32. Chiruvella, K.K., Liang, Z. and Wilson, T.E. (2013) Repair of double-strand breaks by end joining. *Cold Spring Harb Perspect Biol*, **5**, a012757.

33. Sfeir, A. and Symington, L.S. (2015) Microhomology-Mediated End Joining: A Back-up Survival Mechanism or Dedicated Pathway? *Trends Biochem Sci*, **40**, 701-714.

34. Price, B.D. and D'Andrea, A.D. (2013) Chromatin remodeling at DNA double-strand breaks. *Cell*, **152**, 1344-1354.

35. Rogakou, E.P., Pilch, D.R., Orr, A.H., Ivanova, V.S. and Bonner, W.M. (1998) DNA double-stranded breaks induce histone H2AX phosphorylation on serine 139. *The Journal of biological chemistry*, **273**, 5858-5868.

36. Paull, T.T., Rogakou, E.P., Yamazaki, V., Kirchgessner, C.U., Gellert, M. and Bonner, W.M. (2000) A critical role for histone H2AX in recruitment of repair factors to nuclear foci after DNA damage. *Curr Biol*, **10**, 886-895.

37. Mailand, N., Bekker-Jensen, S., Faustrup, H., Melander, F., Bartek, J., Lukas, C. and Lukas, J. (2007) RNF8 ubiquitylates histones at DNA double-strand breaks and promotes assembly of repair proteins. *Cell*, **131**, 887-900.

38. Mattiroli, F., Vissers, J.H., van Dijk, W.J., Ikpa, P., Citterio, E., Vermeulen, W., Marteijn, J.A. and Sixma, T.K. (2012) RNF168 ubiquitinates K13-15 on H2A/H2AX to drive DNA damage signaling. *Cell*, **150**, 1182-1195.

39. Wang, B. and Elledge, S.J. (2007) Ubc13/Rnf8 ubiquitin ligases control foci formation of the Rap80/Abraxas/Brca1/Brcc36 complex in response to DNA damage. *Proc Natl Acad Sci U S A*, **104**, 20759-20763.

40. Kolas, N.K., Chapman, J.R., Nakada, S., Ylanko, J., Chahwan, R., Sweeney, F.D., Panier, S., Mendez, M., Wildenhain, J., Thomson, T.M. *et al.* (2007) Orchestration of the DNA-damage response by the RNF8 ubiquitin ligase. *Science*, **318**, 1637-1640.

41. Huen, M.S., Grant, R., Manke, I., Minn, K., Yu, X., Yaffe, M.B. and Chen, J. (2007) RNF8 transduces the DNA-damage signal via histone ubiquitylation and checkpoint protein assembly. *Cell*, **131**, 901-914.

42. Smeenk, G. and van Attikum, H. (2013) The chromatin response to DNA breaks: leaving a mark on genome integrity. *Annu Rev Biochem*, **82**, 55-80.

43. House, N.C., Koch, M.R. and Freudenreich, C.H. (2014) Chromatin modifications and DNA repair: beyond double-strand breaks. *Front Genet*, **5**, 296.

44. Abu-Farha, M., Lanouette, S., Elisma, F., Tremblay, V., Butson, J., Figeys, D. and Couture, J.F. (2011) Proteomic analyses of the SMYD family interactomes identify HSP90 as a novel target for SMYD2. *J Mol Cell Biol*, **3**, 301-308.

45. Sarris, M.E., Moulos, P., Haroniti, A., Giakountis, A. and Talianidis, I. (2016) Smyd3 Is a Transcriptional Potentiator of Multiple Cancer-Promoting Genes and Required for Liver and Colon Cancer Development. *Cancer Cell*, **29**, 354-366.

46. Zhang, C.Z., Spektor, A., Cornils, H., Francis, J.M., Jackson, E.K., Liu, S., Meyerson, M. and Pellman, D. (2015) Chromothripsis from DNA damage in

micronuclei. *Nature*, **522**, 179-184.

47. Alvarez-Quilon, A., Serrano-Benitez, A., Lieberman, J.A., Quintero, C., Sanchez-Gutierrez, D., Escudero, L.M. and Cortes-Ledesma, F. (2014) ATM specifically mediates repair of double-strand breaks with blocked DNA ends. *Nat Commun*, **5**, 3347.

48. Sulli, G., Di Micco, R. and di Fagagna, F.d.A. (2012) Crosstalk between chromatin state and DNA damage response in cellular senescence and cancer. *Nat Rev Cancer*, **12**, 709-720.

49. Stucki, M. and Jackson, S.P. (2004) MDC1/NFBD1: a key regulator of the DNA damage response in higher eukaryotes. *DNA Repair (Amst)*, **3**, 953-957.

50. Minter-Dykhouse, K., Ward, I., Huen, M.S., Chen, J. and Lou, Z. (2008) Distinct versus overlapping functions of MDC1 and 53BP1 in DNA damage response and tumorigenesis. *J Cell Biol*, **181**, 727-735.

51. Zhenkun Lou, Claudia Christiano Silva Chini, Katherine Minter-Dykhouse and Chen, J. (2003) Mediator of DNA damage checkpoint protein 1 regulates BRCA1 localization and phosphorylation in DNA damage checkpoint control. *The Journal of biological chemistry*, **278**, 13599-13602.

52. Wilson, K.A. and Stern, D.F. (2008) NFBD1/MDC1, 53BP1 and BRCA1 have both redundant and unique roles in the ATM pathway. *Cell cycle*, **7**, 3584-3594.

53. Kolas NK, Chapman JR, Nakada S, Y.J., Chahwan R, Sweeney FD, Panier S, Mendez M, Wildenhain J, Thomson TM, Pelletier L *et al.* (2007) Orchestration of the DNA-Damage Response by the RNF8 Ubiquitin Ligase. *Science*, **318(5856)**, 1637-1640.

54. Shi, W., Ma, Z., Willers, H., Akhtar, K., Scott, S.P., Zhang, J., Powell, S. and Zhang, J. (2008) Disassembly of MDC1 foci is controlled by ubiquitin-proteasome-dependent degradation. *The Journal of biological chemistry*, **283**, 31608-31616.

55. Lou, Z., Chini, C.C., Minter-Dykhouse, K. and Chen, J. (2003) Mediator of DNA damage checkpoint protein 1 regulates BRCA1 localization and phosphorylation in DNA damage checkpoint control. *J Biol Chem*, **278**, 13599-13602.

56. Choudhury, A.D., Xu, H. and Baer, R. (2004) Ubiquitination and proteasomal degradation of the BRCA1 tumor suppressor is regulated during cell cycle progression. *The Journal of biological chemistry*, **279**, 33909-33918.

57. Phatnani, H.P. and Greenleaf, A.L. (2006) Phosphorylation and functions of the RNA polymerase II CTD. *Genes Dev*, **20**, 2922-2936.

58. Lee, T.Y., Chang, W.C., Hsu, J.B., Chang, T.H. and Shien, D.M. (2012) GPMiner: an integrated system for mining combinatorial cis-regulatory elements in mammalian gene group. *BMC Genomics*, **13 Suppl 1**, S3.

59. Cock-Rada, A.M., Medjkane, S., Janski, N., Yousfi, N., Perichon, M., Chaussepied, M., Chluba, J., Langsley, G. and Weitzman, J.B. (2012) SMYD3 Promotes Cancer Invasion by Epigenetic Upregulation of the Metalloproteinase MMP-9. *Cancer Research*, **72**, 810-820.

60. Zou, J.N., Wang, S.Z., Yang, J.S., Luo, X.G., Xie, J.H. and Xi, T. (2009) Knockdown of SMYD3 by RNA interference down-regulates c-Met expression and inhibits cells migration and invasion induced by HGF. *Cancer letters*, **280**, 78-85.

61. Karanam, K., Kafri, R., Loewer, A. and Lahav, G. (2012) Quantitative live cell imaging reveals a gradual shift between DNA repair mechanisms and a maximal use of HR in mid S phase. *Mol Cell*, **47**, 320-329.

62. Helleday, T. (2010) Homologous recombination in cancer development, treatment and development of drug resistance. *Carcinogenesis*, **31**, 955-960.

63. Ren, T.N., Wang, J.S., He, Y.M., Xu, C.L., Wang, S.Z. and Xi, T. (2011) Effects of SMYD3 over-expression on cell cycle acceleration and cell proliferation in MDA-MB-231 human breast cancer cells. *Med Oncol*, **28 Suppl 1**, S91-98.

64. Liu, C., Wang, C., Wang, K., Liu, L., Shen, Q., Yan, K., Sun, X., Chen, J., Liu, J., Ren, H. *et al.* (2013) SMYD3 as an oncogenic driver in prostate cancer by stimulation of androgen receptor transcription. *J Natl Cancer Inst*, **105**, 1719-1728.

65. Vieira, F.Q., Costa-Pinheiro, P., Almeida-Rios, D., Graca, I., Monteiro-Reis, S., Simoes-Sousa, S., Carneiro, I., Sousa, E.J., Godinho, M.I., Baltazar, F. *et al.* (2015) SMYD3 contributes to a more aggressive phenotype of prostate cancer and targets Cyclin D2 through H4K20me3. *Oncotarget*, **6**, 13644-13657.

66. Piao, L., Kang, D., Suzuki, T., Masuda, A., Dohmae, N., Nakamura, Y. and Hamamoto, R. (2014) The histone methyltransferase SMYD2 methylates PARP1 and promotes poly(ADP-ribosyl)ation activity in cancer cells. *Neoplasia*, **16**, 257-264, 264 e252.

67. Jha, D.K. and Strahl, B.D. (2014) An RNA polymerase II-coupled function for histone H3K36 methylation in checkpoint activation and DSB repair. *Nat Commun*, **5**, 3965.

68. Pfister, S.X., Ahrabi, S., Zalmas, L.P., Sarkar, S., Aymard, F., Bachrati, C.Z., Helleday, T., Legube, G., La Thangue, N.B., Porter, A.C. *et al.* (2014) SETD2-dependent histone H3K36 trimethylation is required for homologous recombination repair and genome stability. *Cell Rep*, **7**, 2006-2018.

69. Aymard, F., Bugler, B., Schmidt, C.K., Guillou, E., Caron, P., Briois, S., Iacovoni, J.S., Daburon, V., Miller, K.M., Jackson, S.P. *et al.* (2014) Transcriptionally active chromatin recruits homologous recombination at DNA double-strand breaks. *Nat Struct Mol Biol*, **21**, 366-374.

70. Kim, J.M., Kim, K., Schmidt, T., Punj, V., Tucker, H., Rice, J.C., Ulmer, T.S. and An, W. (2015) Cooperation between SMYD3 and PC4 drives a distinct transcriptional program in cancer cells. *Nucleic acids research*, **43**, 8868-8883.

71. Ben-Porath, I., Thomson, M.W., Carey, V.J., Ge, R., Bell, G.W., Regev, A. and Weinberg, R.A. (2008) An embryonic stem cell-like gene expression signature in poorly differentiated aggressive human tumors. *Nat Genet*, **40**, 499-507.

72. Li, H. and Durbin, R. (2010) Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics*, **26**, 589-595.

73. Langmead, B., Trapnell, C., Pop, M. and Salzberg, S.L. (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol*, **10**, R25.

74. Zhu, L.J., Gazin, C., Lawson, N.D., Pages, H., Lin, S.M., Lapointe, D.S. and Green, M.R. (2010) ChIPpeakAnno: a Bioconductor package to annotate ChIP-seq and ChIP-chip data. *BMC Bioinformatics*, **11**, 237.

75. Huang da, W., Sherman, B.T. and Lempicki, R.A. (2009) Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large

gene lists. *Nucleic Acids Res*, **37**, 1-13.

76. Huang da, W., Sherman, B.T. and Lempicki, R.A. (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc*, **4**, 44-57.

77. Olive, P.L. and Banath, J.P. (2006) The comet assay: a method to measure DNA damage in individual cells. *Nat Protoc*, **1**, 23-29.

78. Untergasser, A., Nijveen, H., Rao, X., Bisseling, T., Geurts, R. and Leunissen, J.A. (2007) Primer3Plus, an enhanced web interface to Primer3. *Nucleic Acids Res*, **35**, W71-74.

79. Pierce, A.J., Johnson, R.D., Thompson, L.H. and Jasin, M. (1999) XRCC3 promotes homology-directed repair of DNA damage in mammalian cells. *Genes Dev*, **13**, 2633-2638.

80. Tseng, S.F., Chang, C.Y., Wu, K.J. and Teng, S.C. (2005) Importin KPNA2 is required for proper nuclear localization and multiple functions of NBS1. *J Biol Chem*, **280**, 39594-39600.

81. Kostyrko, K. and Mermod, N. (2016) Assays for DNA double-strand break repair by microhomology-based end-joining repair mechanisms. *Nucleic Acids Res*, **44**, e56.
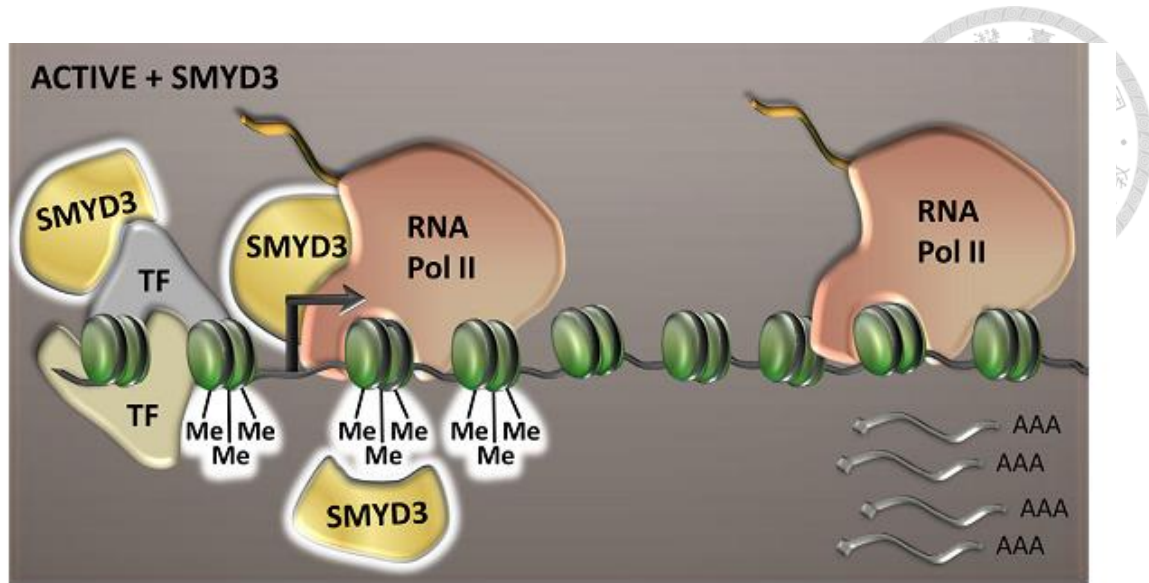
Chapter 8. Appendix.

Figures Contributed by Other Authors.

Appendix Figure 1 were a picture adapted from Antonis Giakountis *et al*., 2017 (13).

Appendix Figure 2, 4 and 6 were performed by Pin-Yu Wang.

Appendix Figure 3 and 5 were performed by Cheng-Hui Tsai.

**Appendix Figure 1. Function of nuclear SMYD3.**

SMYD3 acts as a potentiator of oncogenic transcription associated with cell proliferation, cell cycle activation, angiogenesis and epithelial-to-mesenchymal transition. SMYD3 is localized in the nucleus and potentiates oncogenic transcription through its recruitment to already active gene regulatory regions via interactions with H3K4me3, RNA Pol II or transcription factors (TF).

63
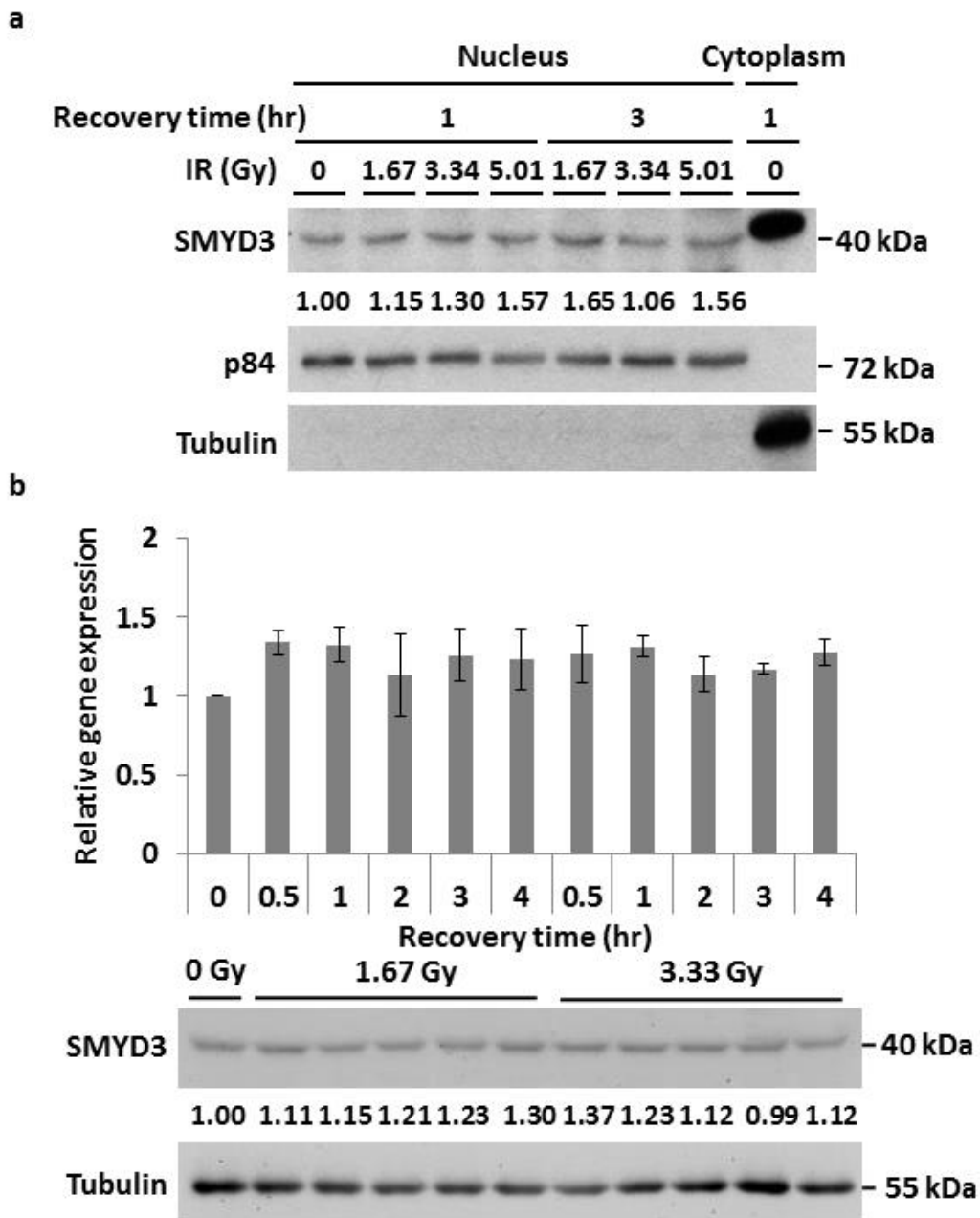
**Appendix Figure 2. SMYD3 location and expression are not modulated by DNA damage insults.**

(a) The amounts of SMYD3 in the nucleus was analyzed by nuclear/cytosol fractionation at 1 hr and 3 hr post-IR with different dosages. Western blot analyses were performed using indicated antibodies. p84 was used as a nuclear-specific marker for the

64

loading control. Tubulin was used to check the purity of nuclear fractions. **(b)** SMYD3

expression levels at 1 hr and 3 hr post-IR with different dosages. Upper panel, the

mRNA expression levels *of SMYD3* were examined by qRT-PCR. Lower panel, the

protein expression levels of SMYD3 were detected by Western blot analyses. All values

in the histograms were means ± SD of triplicates and data were representative of n ≥ 3

for each experiment.
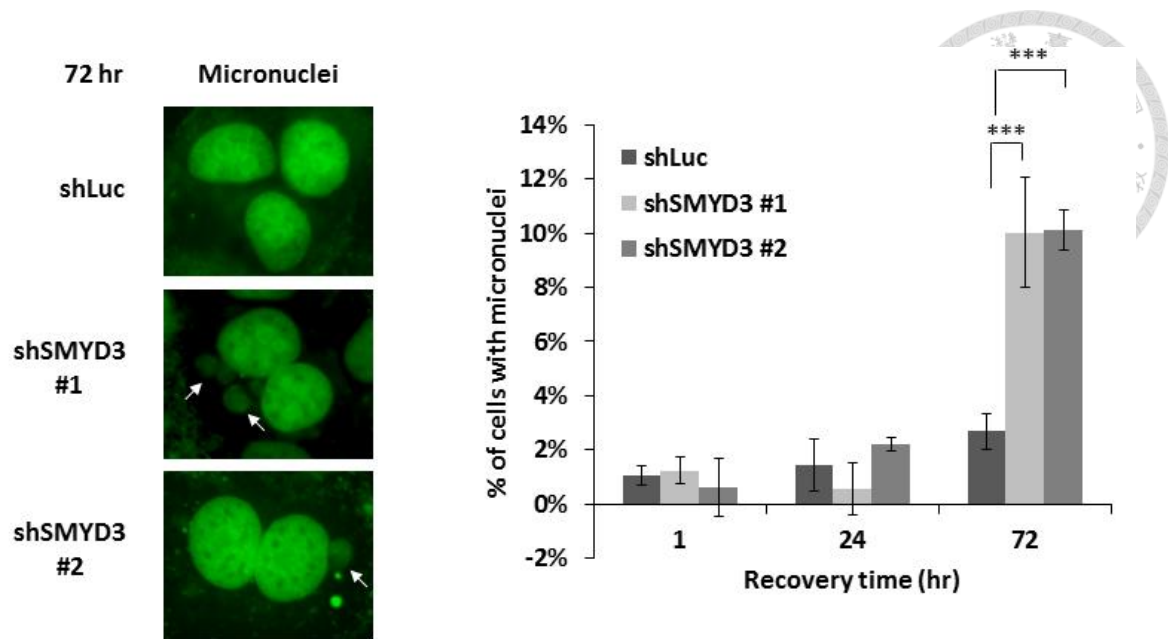
Appendix Figure 3. SMYD3 deficiency increases the ratio of micronuclei after IR treatment.

Right panel, representative images of micronuclei in shLuc and shSMYD3 MCF7 cells treated with 1.67 Gy IR and recovered for 72 h. White arrows point to micronuclei, which signify aberrant chromosomal segregation. Left panel, quantification of micronuclei at indicated times. ***, $P < 0.001$. A minimum of 100 cells per treatment group were analyzed for each quantification. All values in the diagrams were means $\pm$ SD of triplicates and data were representative of $n \geq 3$ for each experiment.

66
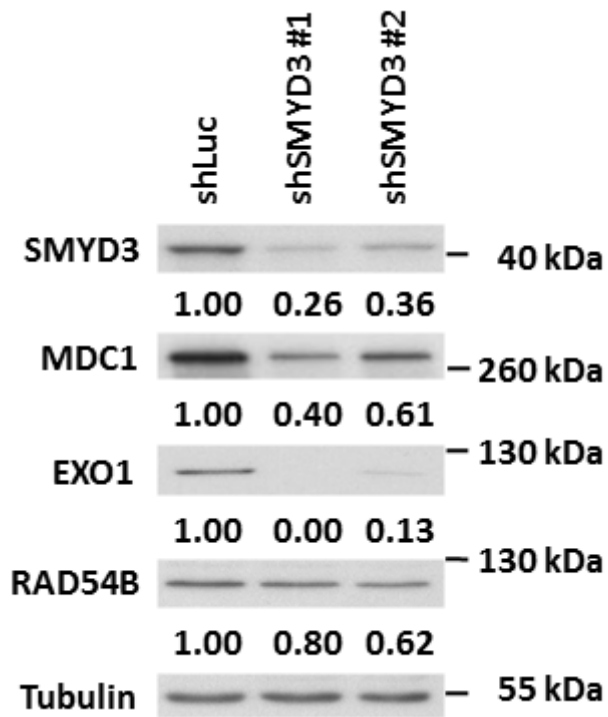
Appendix Figure 4. SMYD3 knockdown downregulates HR gene expressions.

Western blot analyses of SMYD3, MDC1, EXO1 and RAD54B in shLuc and shSMYD3 MCF7 cells. Tubulin was used as a loading control. The ratio of individual protein relative to the loading control tubulin was indicated below.
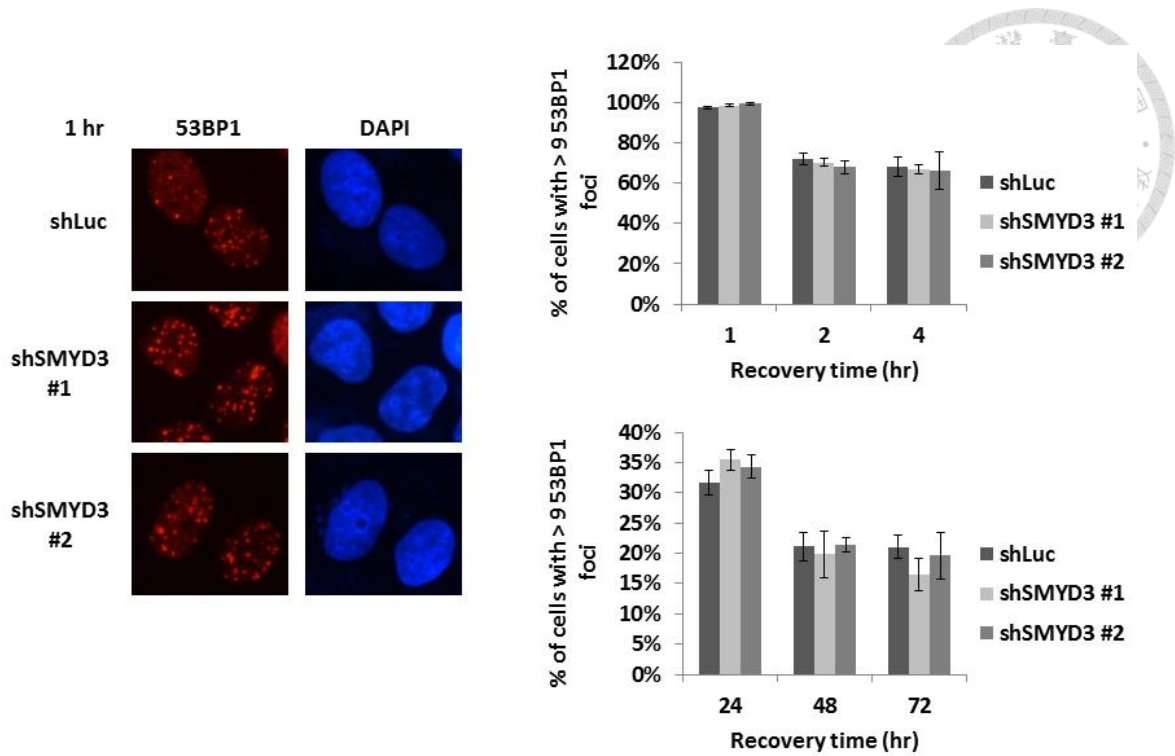
Appendix Figure 5. SMYD3 deficiency does not affect the formation of 53BP1 foci after IR treatment.

The formation of 53BP1 foci at indicated times after 1.67 Gy IR treatment in shLuc or shSMYD3 MCF7 cells. *, $P < 0.05$. **, $P < 0.01$. ***, $P < 0.001$ vs. shLuc control. All values in the histograms *were* means ± SD of *triplicates* and data were representative of n ≥ 3 for each experiment.

68
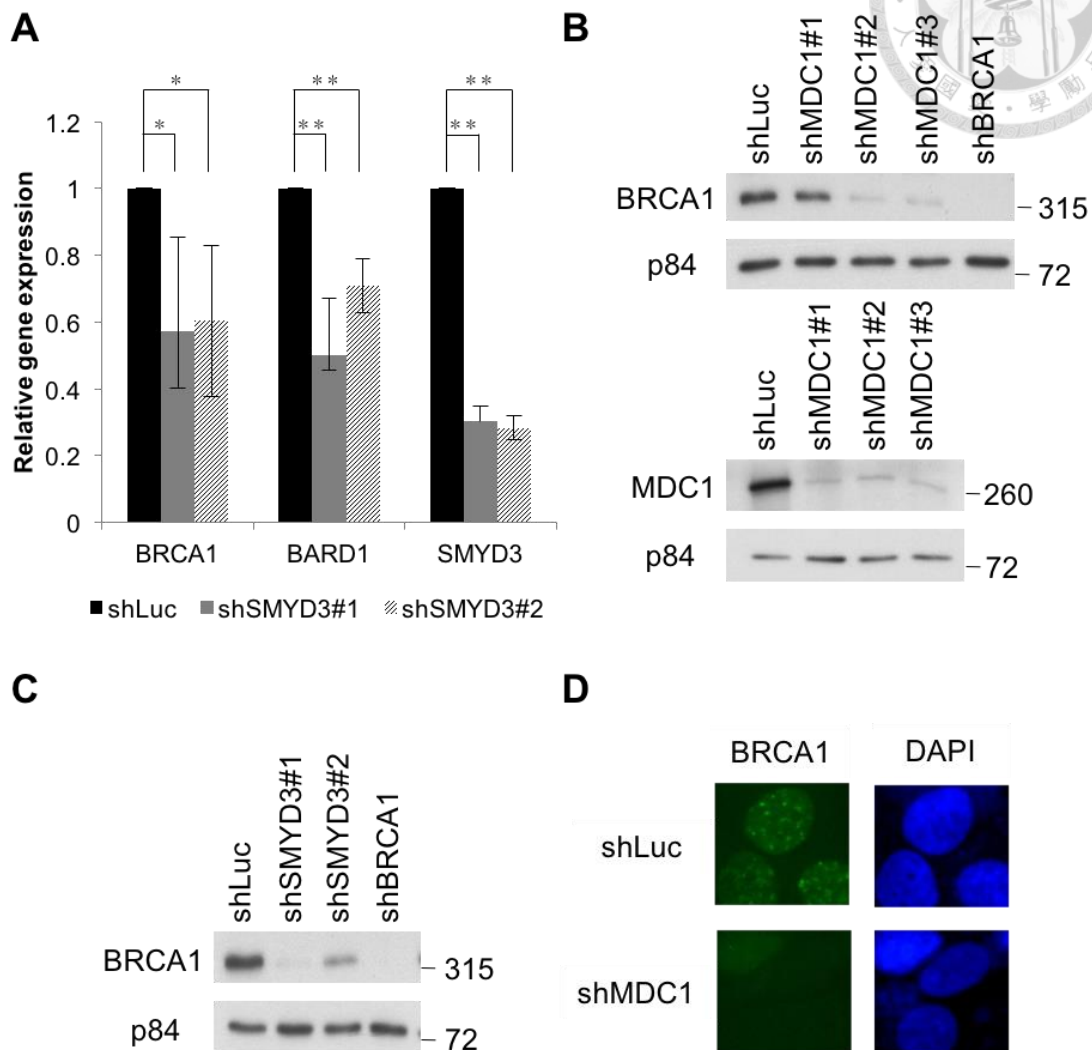
Appendix Figure 6. BRCA1 is reduced in SMYD3-depleted cells.

(A) Expression of BRCA1 and BARD1 decreased in SMYD3 knockdown cells. Error

bar represents the mean $\pm$ S.D. *, $P < 0.05$. **, $P < 0.01$. (B) BRCA1 reduced in three

knockdown clones of MDC1. Knockdown efficiency of MDC1 was shown. p84 was

used as internal control. (C) BRCA1 reduced in SMYD3-depleted cells. (D)

Representative micrographs of BRCA1 in MDC1-depleted cells after IR treatment.

69